Universiteit
Leiden
The Netherlands

Mind the gap : predicting cardiovascular risk during drug development

Chain, A.S.Y.

Citation

Chain, A. S. Y. (2012, October 9). *Mind the gap : predicting cardiovascular risk during drug development*. Retrieved from https://hdl.handle.net/1887/19939

Version:	Corrected Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/19939

Note: To cite this publication please use the final published version (if applicable).

Cover Page

Universiteit Leiden

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

Anne S. Y. Chain

The research presented in this thesis was performed within the framework of project D2-104 of the Dutch Top Institute Pharma at the Division of Pharmacology of the Leiden-Amsterdam Center for Drug Research (LACDR), Leiden University, Leiden, The Netherlands

Cover design by Omar Duragos, Creative Artist
Nomad Creative, Toronto, Canada
Nomad.Creative@rogers.com

Printing by Ipskampdrukkers,
Enschede, The Netherlands

ISBN: 978-94-6191-416-3

Copyright ©2012 by Anne S. Y. Chain

All rights reserved. No part of this thesis may be reproduced, distributed, stored in a retrieval system, or transmitted in any form or by any means, without the written permission of the author and the publisher holding the copyright of the published articles.

Universiteit Leiden

MIND THE GAP

PREDICTING CARDIOVASCULAR RISK DURING DRUG DEVELOPMENT

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. P. F. van der Heijden,
volgens besluit van het College voor Promoties
te verdedigen op dinsdag 9 oktober 2012
klokke 15.00 uur

door

ANNE S. Y. CHAIN
geboren te Hong Kong
in 1982

Promotiecommissie:

Promotoren:	*Prof. dr. M. Danhof* *Prof. dr. M.C.J.M. Sturkenboom, Eramus Universiteit Rotterdam*
Co-promotor:	*Dr. O.E. Della Pasqua*
Leden:	*Prof. dr. L. Aarons, The University of Manchester, Manchester, UK.* *Prof. dr. J. Burggraaf, CHDR, Leiden* *Prof dr. A.C.G. Egberts, Universiteit van Utrecht, Utrecht* *Prof dr. T. Hankemeier, Universiteit Leiden, Leiden* *Prof dr. A.P. IJzerman, Universiteit Leiden, Leiden*

To Nai-Nai and Yeh-Yeh

wishing you were here.....

CONTENTS

SECTION I	**GENERAL INTRODUCTION**	**11**
Chapter 1	An introduction to Mind the Gap: predicting cardiovascular risk during drug development	13
Chapter 2	Scope and intent of investigation	59
SECTION II	**BAYESIAN METHODOLOGY**	**75**
Chapter 3	Assessing the probability of QTc-interval prolongation in clinical drug development	77
Chapter 4	Translational pharmacology of QTc-interval prolongation	103
SECTION III	**FROM PHARMACOLOGY TO EPIDEMIOLOGY - A BRIDGING STRATEGY FOR RISK MANAGEMENT**	**137**
Chapter 5	Not-in-Trial Simulation I: bridging cardiovascular risk from clinical trials to real life conditions	139
Chapter 6	Not-In-Trial Simulation II: predicting cardiovascular risk from clinical trial data	159
Chapter 7	Validation of Not-In-Trial-Simulation: predicting cardiovascular risk from clinical trial data	181
Chapter 8	Can QTc prolongation >10 msec predict sudden cardiac death?	201

SECTION IV **CLINICAL TRIAL SIMULATION IN THE EVALUATION OF CARDIOVASCULAR SAFETY** **219**

Chapter 9 Model-Based Simulation of QT and RR Intervals for the Evaluation of QTc-interval Prolongation 221

Chapter 10 Model-Based Analysis of Thorough QT Studies 237

Chapter 11 Can First-Time-In-Human Trials Replace Thorough QT Studies? 257

SECTION V **CONCLUSIONS & PERSPECTIVE** **281**

Chapter 12 Mind the Gap: Predicting Cardiovascular Risk During Drug Development - Summary, Conclusion and Perspectives 283

Chapter 13 Nederlandse Samenvatting (Synopsis in Dutch) 307

EPILOGUE **331**

Acknowledgement i

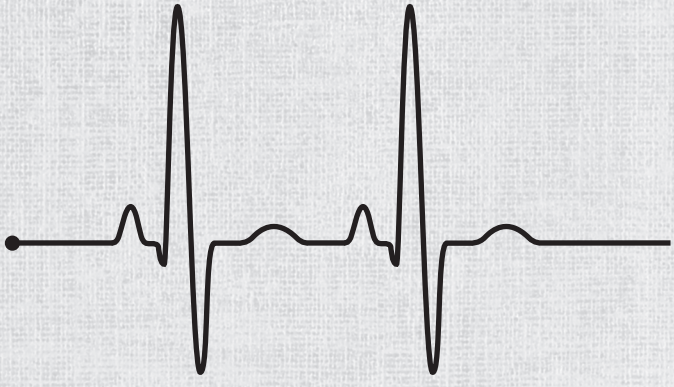
PhD Portfolio iv

List of Publications viii

About the Author iv

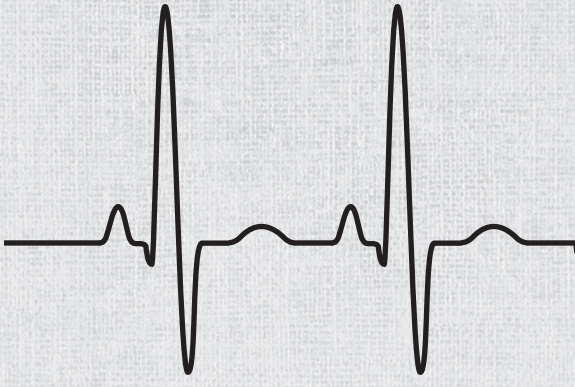
List of Abbreviations x

Stellingen / Propositions to this dissertation xii



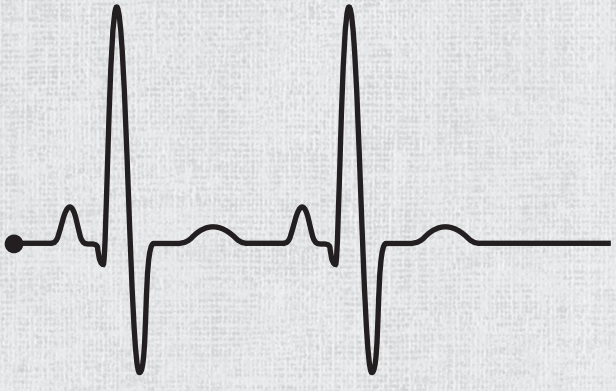
"Devising and testing an effective risk assessment and risk management strategy is where real value for patient is created."

Sir Colin Dollery, EACPT 2009



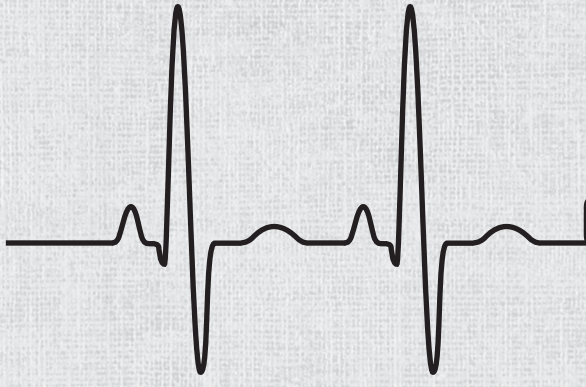
Section I

General Introduction



An Introduction To Mind The Gap:
Predicting Cardiovascular Risk During Drug Development

Published with adaptations in Drug Discovery Today - Technology - in press



Chapter 1

1.1 BACKGROUND

Safety issues remain the major cause of drug attrition during clinical development as well as of post-market drug withdrawals [1-5], accounting for 15-34% of all drug discontinuation [6, 7]. In the last decade cardiovascular safety, specifically drug-induced QT/QTc-interval prolongation, has been an increasing concern for both public health and in drug development [1]. Historically, however, the earliest well documented cases of QT-interval prolongation date back to 1943 regarding the use of quinidine [8, 9]. Quinidine was first introduced as an antiarrhythmic drug in 1918 [10]. Its antiarrhythmic efficacy was attributed to its ability to prolong cardiac repolarisation duration. On the other hand, its use was soon followed by reports of syncopal attacks and unexpected fatalities, and these adverse episodes were attributed to respiratory paralysis or ventricular irritability [11-13]. It was considered that QT-interval prolongation was an early significant indication of the onset of its activity. In 1958, thioridazine was introduced to the market for the treatment of schizophrenia. The observations that apparently, well-tested non-cardiac drugs could also have unintended adverse cardiac effects were made as early as 1963. Following further reports, this effect was shown to be caused by electrophysiological changes that reflect on the surface electrocardiogram (ECG) as QT-interval prolongation [14-20]. Concern later extended to other drugs of its chemical class (phenothiazines) [21-27]. By early 1970s, the concern had extended to the entire therapeutic class of antipsychotic drugs.

Given the potentially fatal consequences of this concentration-dependent adverse drug reaction, regulatory authorities have reacted to this relatively recent "pharmacoepidemic" by denying or delaying the approval of a number of new drugs and placing severe restrictions on the use of many old and some new drugs because of concerns arising from their potential to prolong the QT/QTc-interval.

In the subsequent paragraphs of this chapter, we provide an overview of the physiological, clinical and epidemiological basis for the characterisation of drug-induced effects on cardiac conductivity, more specifically on QTc-interval prolongation. A summary of the mainstream techniques used in *in vitro* and *in vivo* studies is presented, which provides insight into some of the challenges for the translation and prediction of drug effects in humans. 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. Lastly, we introduce important methodological concepts pertinent to the evaluation of exposure-response relationships and the utility of these relations for the prediction of safety of compounds in drug development.

1.2 ELECTROPHYSIOLOGY OF VENTRICULAR ARRHYTHMIAS

QT-interval of in the surface electrocardiogram (ECG), typically expressed in milliseconds, represents the ventricular action potential time duration in the heart's sinus rhythm (Figure 1). It is determined by a balance between a number of inward and outward ionic currents, such as sodium (Na^+), potassium (K^+), calcium (Ca^{2+}) and chloride (Cl^-). The primary current is the rapid component of the outward repolarising current (I_{kr}) that is mediated primarily by the delayed rectifying potassium channels.

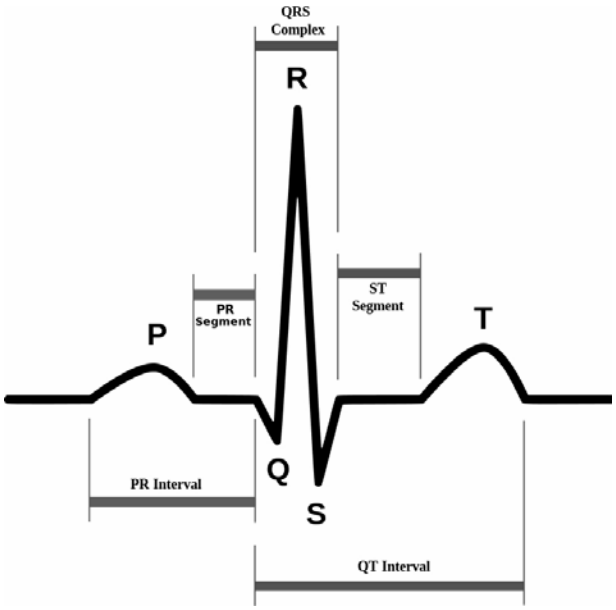


Figure 1. Sinus rhythm of a single heartbeat shown in a surface electrocardiogram (ECG). There are typically six distinct waves (identified by the letters P, Q, R, S, T, and U) and they occur in a specific order, over specific periods of time, with specific relative sizes. While there is a significant range within which variations in rhythm are considered normal, anything that significantly deviates from sinus rhythm by more than a certain amount may be indicative of heart disease.
http://en.wikipedia.org/wiki/File:SinusRhythmLabels.svg, as accessed on 27 March, 2012 at 15:44 EST.

Decreases in ionic currents leading to the prolongation of the QT-interval can be a result of electrolyte imbalance, endocrine dysfunction, autonomic imbalance, various disease states or following administration of drugs (as an intended effect or not). Drug-induced prolongation is expected with class III antiarrhythmic drugs (potassium channel blockers) which are intended to produce their desired therapeutic benefit by blocking I_{kr}, delaying ventricular repolarisation and thereby increasing the myocardial refractory period. Some typical drugs in this category include *d,l*-sotalol, ibutilide, dofetilide, terikalant, azimilide and N-acetyl-procainamide.

The QT-interval is also prolonged when the I_{kr} current is diminished as a result of genetic mutations, such as the Romano-Ward syndrome, a major variant of long-QT syndrome (LQTS) [28]. All forms of LQTS will involve an abnormal repolarisation of the heart causing differences in the refractory period of the heart muscle cells. As a result of an increase in the frequency of the prolonged repolarisation, early after-depolarisations (EADs) can occur, which can degrade to various types of arrhythmias such as torsade de pointes (TdP), ventricular fibrillation and tachycardia that can eventually lead to sudden cardiac death (SCD) (Figure 2). By contrast, the incidence of sudden death in young people without structural heart disease is approximately 1 in 2500 [29-31].

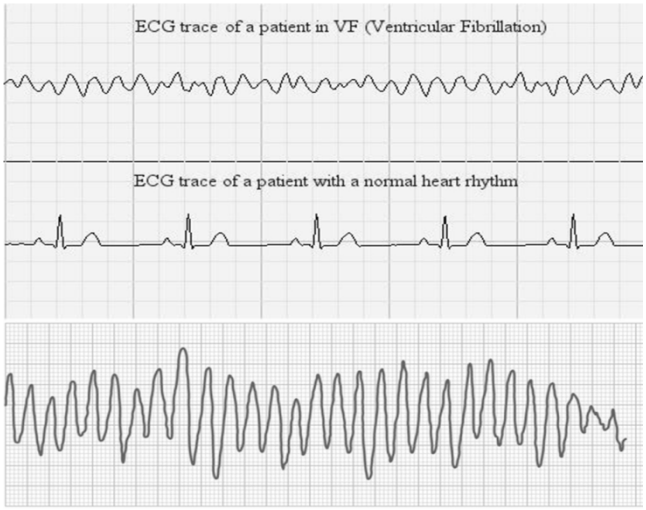


Figure 2. Electrocardiograms of patient in ventricular fibrillation (top) compared to one with normal heart rhythm (middle) and one with torsade de pointes (bottom). http://www.homeheart.co.uk/html/glossary.html, http://medlibes.com/entry/torsades-de-pointes, as accessed 19 April, 2012 at 9:03 EST.

1. 2. 1 TORSADE DE POINTES

QT-intervals are very much related to heart rate, designated as the RR-intervals (measured as the time interval between consecutive R peaks). The relationship is that the faster the heart rate, the shorter the QT-interval. Thus QT-intervals are often adjusted / corrected for heart rate to describe the true action potential duration in an individual. Typically population-based correction factors are used in clinical practice. Bazett's and Fridericia's formula have become standard methods for routine ECG monitoring, even though individual correction factors have been shown to better account for individual differences.

BAZETTs' FORMULA: $\boldsymbol{QTcB=QT/\sqrt{RR}}$ [32]

FRIDERICIA's FORMULA: $\boldsymbol{QTcF=QT/\sqrt[3]{RR}}$ [33]

Excessive RR-corrected QT (QTc)-interval prolongation can be proarrhythmic and degenerate into a potentially fatal ventricular tachyarrhythmia known as torsade de pointes (TdP), a unique polymorphic form of ventricular tachycardia which is associated with concomitant prolonged QTc-interval [34] (Figure 3). The European regulatory guideline, based on the opinion of an ad-hoc expert group, used three gender-specific categories to differentiate QTc prolongation [35]. For women, normal is defined as $\leq$ 450 msec, 451-470 msec as borderline and > 470 msec as prolonged. For men, $\leq$ 430 msec is normal, 431 to 450 msec is designated as borderline and prolonged is used when QTc-intervals are > 450 msec.

Although QTc-interval prolongation is not necessarily harmful, it is at present the best and only surrogate marker for TdP, especially so because the definition of this particular arrhythmia requires prolongation of QT-interval as a preceding event. One review in 1993 concluded "at present, our knowledge base about the relation of the QT-interval and TdP is grossly incomplete" [34]. Unfortunately, despite extensive research since then, this still remains the case today. To a large extent, the (near) theological reliance on QT-interval prolongation as a surrogate of TdP originates from the very definition of this unique polymorphic ventricular tachyarrhythmia [34]. Ventricular tachyarrhythmias, even when meeting the morphological criteria of TdP are not labelled by many clinicians as TdP unless preceded by QTc-interval prolongation [37-39].

TdP is triggered by the appearance of early after-depolarisation (EADs), mediated by slow inward calcium current, during the late phase 2 of the prolonged cardiac action potential. Therefore as an extension of their pharmacological effect, this iatrogenic proarrhythmia may be expected in some individuals following the use of antiarrhythmic drugs, which possess potassium channel blocking activity. Critically, TdP is a potentially fatal adverse drug reaction. That is because TdP subsequently degenerates into ventricular fibrillation in about 20% of cases [40] and not uncommonly, leads to cardiac arrest and even sudden death may be the outcome [41]. The overall mortality is of the order of 10-17% [40, 42].

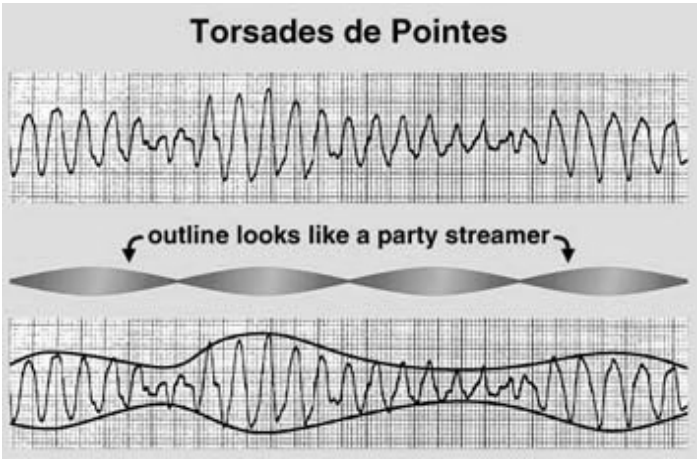


Figure 3. Electrocardiogram of torsade de pointes. http://www.doctortipster.com/3748-ventricular-tachycardia-causes-diagnosis-and-treatment.html, as access on 19 April, 2012 at 9:12 EST.

The balance between the risk of a fatal adverse event and therapeutic benefit is a very delicate one, depending not only on the drug concerned and its plasma concentration but also on a number of host modifying factors. These risk factors include female gender, electrolyte imbalance (especially hypokalaemia), myocardial ischaemia, congestive heart failure, bradycardia and pre-existing prolongation (i.e. long-QT syndrome), which all result in inherently prolonged baseline QTc values. Moreover, risk to an individual can vary from day to day. An inter-current event or therapeutic intervention for an unrelated condition, such as interacting drugs, can easily increase or put risk on patient who was previously not at risk.

Unfortunately, the potential to prolong the QTc-interval and induce TdP is not confined to class III antiarrhythmia drugs. A number of class I antiarrhythmia and anti-angina drugs as well as non-cardiovascular drugs also carry this liability. In a survey of 2194 cases of TdP in the US Food and Drug Administration (FDA) database [42], the most common drugs implicated belong to the class of cardiac (26.2%), central nervous system (CNS) (21.9%), anti-infective (19.0%) and antihistamine (11.6%) compounds. Of the 2194 cases, 61.1% were associated with hospitalisation, 27.9% were life-threatening and 9.8% were associated with a fatal outcome. The regulatory focus on QTc-interval prolongation by drugs has therefore changed from one of a potentially desirable antiarrhythmia mechanism to one of potentially fatal proclivity.

1. 2. 2 FREQUENCY OF TORSADE DE POINTES

The frequency of TdP, or prolongation of the QTc-interval to proarrhythmic threshold, varies with the class of drugs. It is, not unexpectedly, the highest with class III drugs. For non-cardiac drugs, the frequency is unknown and can vary from approximately 1 in 100 (halofantrine) to 1 in 50 000 (terfenadine), depending on clinical circumstances. However, the true frequency of this effect with non-cardiac drugs is difficult to estimate because the diagnosis of this toxicity requires an ECG-monitoring facility, which is either not available in a general practitioner's surgery or not readily utilised in local hospitals. Due to the fact that TdP can be transient, its diagnosis in a patient presenting with dizziness or syncope requires an immediate access to cardiac rhythm recording facility. Even in asymptomatic patients, despite the requirements included in prescribing information, there is a general lack of appropriate patient monitoring by ECG. More importantly, however, the effect is often not recognised as iatrogenic and it is grossly under-reported (reporting rate in the order to 10-20%) even when recognised as drug-induced. This was exemplified by the events preceding the withdrawal of terodiline [43]. In all likelihood, the frequency of these events is relatively low (<0.1%) and below what could be confidently detected by the size of the clinical trials database included in a regulatory submission. The frequency is sufficiently low, that the risk was uncovered only through spontaneous reports during the post-marketing use of the drugs concerned. Although low, the risk is nonetheless unacceptable for many conditions which are relatively benign in nature. Halofantrine and

arsenic trioxide best illustrate the careful need to balance the potential risk against potential benefits.

Interestingly, 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 individual with equivalent prolongation of the QTc-interval. Neither do all drugs that prolong the QT-interval to an equivalent duration carry the same risk of inducing TdP [44, 45]. The incidence of TdP is estimated to be 0.5-8.8% with quinidine [46] and 2.6 – 4.1% with *d,l*-sotalol [42, 47, 48] (Figure 4). Since QT-interval prolongation is not an ideal surrogate for risk of TdP, withdrawals of drugs that simply prolong the QT-interval reflect a pragmatic and conservative approach to risk management. When found to induce TdP frequently, drugs such as prenylamine, terodiline, terfenadine, astemizole, cisapride and levacetylmethadol were all withdrawn from the market when their benefit/risk ratio was determined to be adverse and safer alternatives were available. On the other hand, arsenic trioxide was approved and is still on the market despite a known high frequency of TdP associated with its use. Likewise, pimozide and thioridazine continue to be available. The re-introduction of sertindole has also been discussed because of its low potential for extra pyramidal adverse effects and lack of reports of TdP despite marked prolongation in QT-interval.

1.3 IMPACT ON DRUG DEVELOPMENT

Based on the aforementioned, regulatory and clinical expectations have evolved towards a pre-approval policy for new chemical entities (NCEs), which is aimed at assessing the potential risk of TdP for any new medicine, from early discovery throughout clinical development.

Different actions have been taken to improve drug screening and selection of compounds. Non-clinical protocols have been introduced to support the selection of the most appropriate NCEs to bring them forward to human trials. Experiments have been devised with the primary objective of eliminating torsadogenic compounds and assessing the propensity of QT-

interval prolongation. Although primary efforts have focused on class effects and other QSAR molecular structural relationships, the search for predictive and translational models for QTc-interval prolongation has remained empirical. The requirements for establishing *in vitro*, *in vivo* and ultimately clinical correlations in a quantitative manner are often ignored despite the relevance of concentration-effect relationships as basis for characterising drug properties *in vivo.*

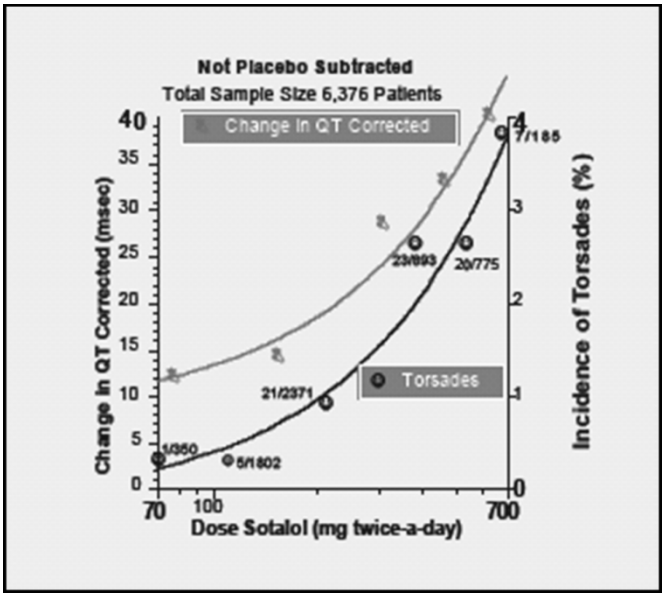


Figure 4. Correlation showing the change in QTc-interval with sotalol dosages and incidence of torsade de pointes [48].

Given the implications for public health, scientific efforts have faced a parallel movement driven by health authorities, which have imposed the introduction of supposedly effective measures for the approval of novel compounds. The EU Committee for Proprietary Medicinal Products (CPMP), known as the Committee for Medicinal Products for Human Use (CHMP) was the first scientific advisory body of a regulatory authority (European Medicines Agency) to issue, in December 1997, a formal guidance note on a strategy by which all NCEs should be investigated for their effect on the QTc-interval [35]. This guidance included recommendations on a set of non-clinical as well as clinical investigations. All strategies devised subsequently are an elaboration of, or minor variations on, the broad pattern set by the CPMP.

In November 2002, the US FDA and Health Canada issued a joint document that focused exclusively on clinical strategies for evaluating the effects of NCEs on QT/QTc-interval prolongation (a preliminary concept paper for discussion) [49]. Following a number of amendments, it was entered into the International Conference on Harmonisation (ICH) process in February 2003 for adoption as a topic that merited global harmonisation, topic E14. The ICH is composed of representatives from regulatory authorities and industry associations in the US, EU and Japan. Representatives from the World Health Organisation (WHO), European Free Trade Area and Canada also attend as observers. In 2005, the guidance introduced the requirement to perform thorough-QT (TQT) studies as the basis for demonstrating a compound's liability to cause QTc prolongation. In addition to outlining the assessment procedures for evaluating prolonged ventricular repolarisation, the ICH E14 document defines a 10 msec increase in QTc-interval as a critical threshold for cardiovascular safety and required the use of a positive control and supra-therapeutic doses of the investigational drug to ensure accuracy and sensitivity of the experimental protocol [50]. Suggestions are also given regarding the timing of the studies as well as the methodologies and interpretations used in the evaluation of QT measurements.

In parallel, the ICH S7B guideline was also created to deal with specific aspects of the non-clinical evaluation of the potential for delayed ventricular repolarisation by human pharmaceuticals. In brief, it promotes integrated risk assessment, based on the chemical and pharmacological class of the drug together with data from two core tests – *in vitro* I_{kr} or hERG assays and *in vivo* studies in a suitable species. Much progress has been made in the last 10 years, but no single non-clinical assay has an absolute positive and/or negative predictive value nor any can be considered a gold standard. Therefore the use of several *in vitro* assays and/or *in vivo* models together, contributes to accurate decision-making and is recommended by most experts in the field.

From a clinical development perspective, translational challenges still exist despite the aforementioned regulatory considerations. Among other issues, insight is lacking into the impact of interacting factors, such as comorbidities, on the actual cardiovascular risk in the target population. Additionally an implicit assumption is made about the magnitude of the drug-induced effects in the population of interest and of the sensitivity of the patient population.

In the next sections, we will describe some key elements underpinning the efforts to translate, predict and bridge data across different phases of drug development. We will show that the challenges to be overcome are partly due to the fact that each stage of the drug development process remains largely modular rather than integrative, with empirical evidence as the basis for developing criteria and decision making on safety matters.

1.4 TRANSLATIONAL EFFORTS: NON-CLINICAL TO CLINICAL

A summary of the mainstream techniques and experimental conditions used in *in vitro* and *in vivo* studies is presented below with the aim of highlighting the lack of a common strategy, i.e., the assessment of pharmacokinetic-pharmacodynamic relationships, which could be used as common denominator for the translation of non-clinical findings into clinically relevant measures. This is followed by a discussion on the various methods used to establish correlations between non-clinical studies and human clinical trials. A summary table (Table 1) is also provided at the end of the section highlighting the main points discussed.

1. 4. 1 *IN VITRO* EXPERIMENTS

hERG Assays

Considered a pivotal step in the lead optimisation phase, the hERG assay is an *in vitro* assay designed to investigate the inhibition of the potassium current through a channel encoded by a human ether-a-go-go-related gene (hERG) [51]. An example of such analysis is shown in Figure 5 [52]. The hERG channel is the rapid component of the delayed rectifier potassium current (I_{kr}) which is the key determinant of the duration of cardiac repolarisation, also known as the QT-interval and thus used to assess the proarrhythmic potential of NCEs [53]. QT-interval prolongation, which can potentially, but not necessarily lead to a form of fatal arrhythmia, torsade de pointes (TdP), is found to be strongly linked with blockage of the hERG channel. However, not all hERG-blockers will result in TdP [54, 55]. Several non-cardiovascular drugs have been withdrawn from the market because of their potential to induce QT-interval prolongation in the surface electrocardiogram (ECG) [56]. It has been reported that the ratio of the

hERG / I_{kr} assay IC50 (concentration that inhibits the hERG current by 50%) values to the free Cmax (maximal plasma concentration) observed in humans after administration of a therapeutic dose, is a useful marker of the potential QT-interval prolonging effect [55, 57]. In most cases, the ratio for drugs associated with QT prolongation has been less than 30-fold; however, there are isolated instances of TdP in humans reported for ratios greater than 30-fold [55, 58]. Thus QT-interval prolongation is not always reliably correlated with hERG blockers or the occurrence of TdP [59]. Lack of confidence in the predictive value of the preclinical tests stems from the fact that each drug associated with TdP in human appears to tell a different story in terms of its electrophysiological profile. There is a consensus about the importance of interactions with the hERG channel but experiments are still not consistently conducted in a quantitative manner and there is no agreement on what constitutes a safe margin [55, 60].

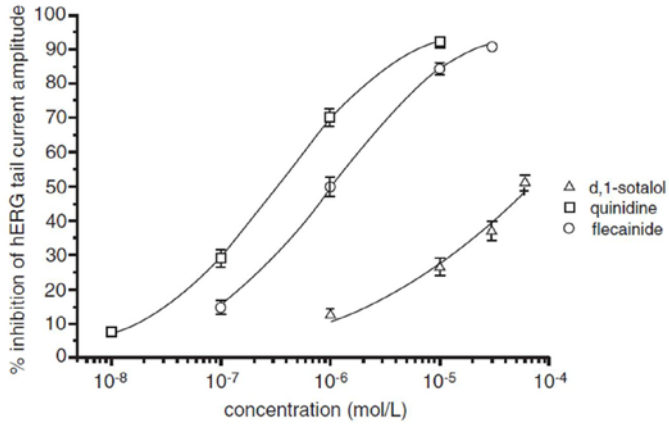


Figure 5. Concentration–response relationship for the inhibitory effects of quinidine, flecainide and d,l-sotalol on hERG tail currents. The estimated IC50 and Hill slope values were, respectively, 0.33 µmol/L and 0.74 for quinidine, 1.03 µmol/L and 0.73 for flecainide and 68.99 µmol/L and 0.50 for d,l-sotalol. Each point with a vertical bar indicates the mean±SEM from six different experiments per compound. Although differences in potency are observed, this system is highly sensitive to experimental conditions. Differences in pH, buffer and voltage level produce differences in the potency and in the concentration-response curve, irrespective of the drug [52].

Action Potential Duration (APD) Assays

The ICH topic S7B safety pharmacology guidance discusses the use of multi-cellular system such as the Purkinje fibres and papillary muscle assay to assess drug-induced changes in repolarisation of action potentials (AP) [61].

Figure 6 summarises the potential mechanisms of APD [62]. These changes provide insight to the mechanisms underlying the genesis of arrhythmias from early after-depolarisation (EAD) making these assays theoretically useful in predicting clinical outcomes [63-65]. Cell preparations for *in vitro* assays are often obtained from different laboratory animal species including rabbit, ferret, guinea pig, dog, mini-pigs and occasionally from humans, given no real industry benchmark has been established [66]. Irrespective of the lack of gold standards, many have exposed the deficiencies in using Purkinje fibres where there is an obvious lack of correlation between the concentration required for hERG IC50 and 10% prolongation in APD90 (the duration at 90% of repolarisation), shown in Figure 7 [57, 67-70].

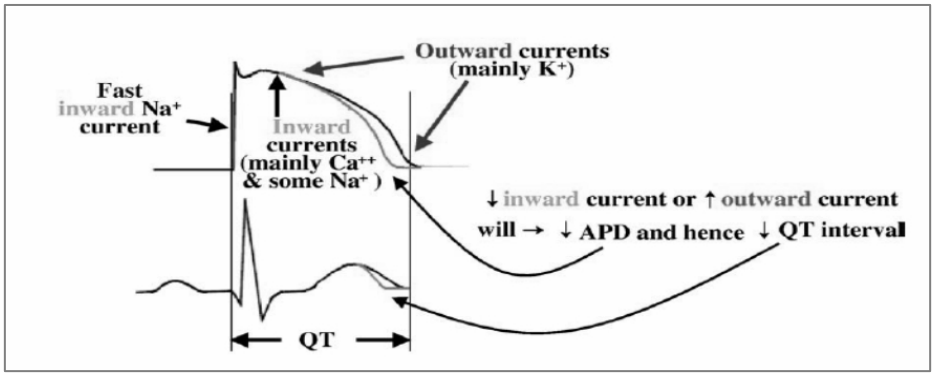


Figure 6. A summary of the potential mechanisms underlying the changes in the action potential duration (APD) [62].

Although the aforementioned non-clinical tests are adequate in predicting the risk of QT-interval prolongation, they are not necessarily sufficient in finding the propensity of pro-arrhythmic risks [59]. Figure 8 demonstrates the significant role of an individual's heart response to proarrhythmic activity as some hearts may develop arrhythmia and others may not in the presence of drug [71]. Thomsen et al. have shown that there is little relationship between the magnitude of the QT-interval and the incidence of TdP [72]. Furthermore, numerous publications have suggested better predictors of TdP in humans [54, 73-75]. Electrophysiological markers

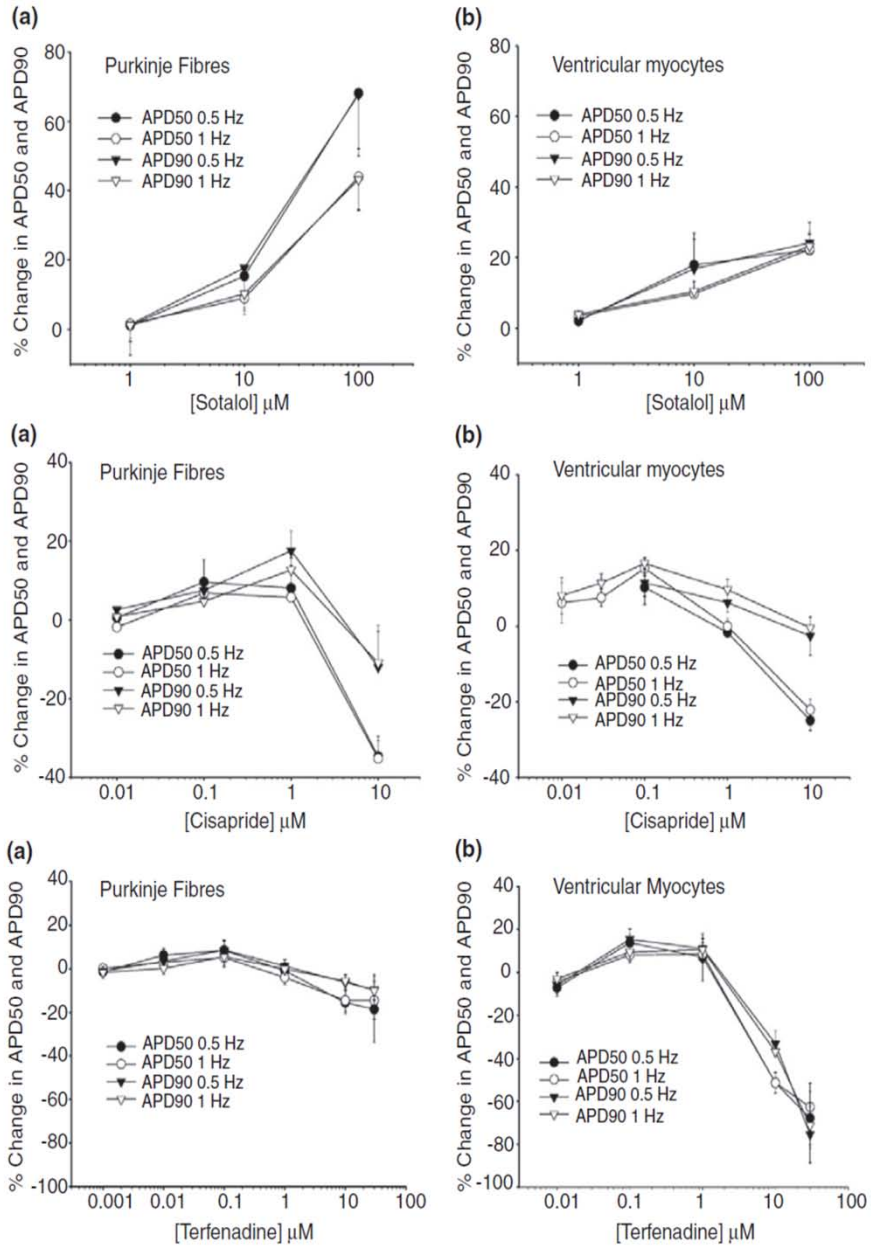


Figure 7. Concentration-dependent effects of d,l-sotalol (top panels), cisapride (middle panels) and terfenadine (lower panels) on action potential duration in dog (Purkinje fibres) and guinea-pig (ventricular myocytes). It is clear that some compounds are sensitive to AP like d,l-sotalol and cisapride when different frequency is used, but flat for terfenadine (insensitive). The difference in the magnitude of effects may dependent on mechanism of action of the compound [70].

associated with drug-induced TdP other than APD and QT-interval include but are not limited to: triangulation; reverse-use dependence; temporal, spatial and transmural dispersion of ventricular repolarisation; the difference in duration between the peak and end of the T wave; and incidence of overt pro-arrhythmias such as EADs and ectopic beats. A version of the Langendorff-perfused female rabbit heart model, (the Screenit system) measures APD, conduction and the TRIaD parameters, namely triangulation, reverse-use dependence, instability and dispersion [76].

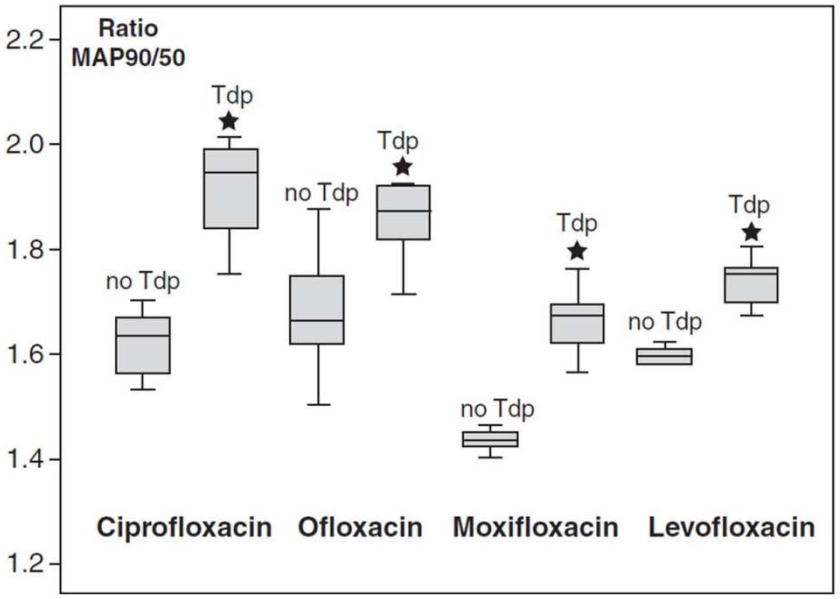


*Figure 8. More triangulation of action potential configuration in hearts is associated with later torsade de pointes (TdP) as compared with hearts without developing TdP *P < 0.05 [71].*

In fact, the Screenit system detects all changes in repolarisation indices and conduction velocity that might interfere with APD values. Changes in the TRIaD parameters have been found to be more predictive than changes in APD for an arrhythmic potential [76]. In several validation studies, Screenit was found to be highly predictive for the pro-arrhythmic potential in drugs. Other species used in the Langendorff heart model include the guinea pig [77-79].

The group of Antzelevitch developed another pro-arrhythmia model based on the transmural heterogeneity in the expression of cardiac ion channels

[73]. Using a left ventricular wedge preparation of the canine heart, trans-membrane AP from epicardian and M-regions were simultaneously measured. The data support the hypothesis that the risk for the development of TdP is related to the increase in transmural dispersion of repolarisation, rather than to prolongation of the QT-interval. This model has been largely superseded by the use of arterially perfused rabbit ventricular wedge, which was found to be more sensitive than dog in terms of the torsadogenic potential when considering clinically relevant concentration ranges in human. Validation with positive controls resulted in a successful translation between preclinical and clinical findings [80].

1. 4. 2 *IN VIVO* STUDIES

Anaesthetised Animals Models

Anaesthetised animals are in general not as sensitive as conscious ones because anaesthetics often induce cardiovascular effects [81]. However, for certain compounds, it is necessary to sedate the animals to prevent unwanted events such as seizures. Anaesthetised dogs and guinea pigs enable the measurement of cardiac contractility and the use of higher doses than in models with conscious animals. Using this model, several cardiovascular parameters can be obtained simultaneously at a modest cost in terms of test compound and animals. The goal of the experiments is often to determine the threshold at which cardiovascular events occur without further consideration of clinical applicability. The lack of commonalities between clinical conditions and experimental settings make the translation of the outcome from these experiments rather challenging.

Conscious Animal Models

ICH S7A and S7B guidelines recommend the use of conscious animals for the assessment of non-clinical cardiovascular safety of a NCE before testing in humans [61, 82]. As a state-of-the-art model, the conscious dog, either sling trained or implanted for telemetry, is used [66, 67, 83]. Usually beagle dogs are used, which allow the model to be combined with complete haemodynamic and pharmacokinetic analysis. However, side effects such as vomiting or CNS-mediated sedation or excitation, at higher doses may occur in conscious animals which may represent a limiting factor in the evaluation of supra-therapeutic doses [83, 84]. The results obtained from *in vivo* dog

studies have been proven to be a robust, sensitive predictor for clinical QT prolongation, but the incidence of TdP is rare [85]. Unfortunately, the current guidelines do not provide specific recommendations on study design, which lead to major inconsistencies among studies conducted by different investigators [86]. This in turn can result in difficulties in translating findings downstream to clinical studies.

In addition to ICH S7A and S7B, other worldwide regulatory guidelines for drug safety evaluation recommend testing both in a rodent and a non-rodent species. Non-human primates, which are phylogenetically close to humans, are often thought of as the "ideal" non-rodent species [87]. Historically, because of ethical issues, ethics, three Rs (replacement, reduction, refinement), bio-safety concerns, price and supply issues, the use of non-human primates in toxicology programmes has been restricted to special cases. Recently, however, biopharmaceutical scientists have turned to primates as the only non-human species in which the biological activities of some drugs are expressed. Even though attention should be paid to compounds for which the metabolites and/or metabolic rates differ between humans and monkeys and the HR changes are beyond the correctable range for the QTc_{Bazett} interval, *in vivo* QT assays using telemetry systems in conscious cynomolgus monkeys are considered to be a sensitive and useful model for assessing the potential for drug-induced QT prolongation in humans [88, 89].

In contrast, the use of the mini-pig in pre-clinical research has been steadily growing in recent years, and is set to continue as experience accumulates. Studies performed on minipigs are fully acceptable by regulatory authorities. Mini-pigs offer several advantages as a second and even first non-rodent species in safety assessment. Furthermore, mini-pigs can be used for invasive and non-invasive cardiovascular safety pharmacology studies. Domestic pigs share many of their features with humans, including anatomy, physiology and biochemistry. In particular, the pigs' cardiovascular system, skin and digestive tract are good models for humans, giving reliable predictions of the toxicity of drugs and chemicals. The use of mini-pigs in regulatory safety testing is therefore not a new idea. Mini-pigs provide better prediction of drug safety than traditional rat or dog models, but have advantages that go beyond their biological similarity to humans [90]. Mini-pigs have a manageable size and are relatively sedentary. Furthermore, Markert et al. have demonstrated a dose-dependent QT prolongation when moxifloxacin

was administered orally at doses that produce clinically relevant plasma drug concentrations [91].

Smaller animals, such as rodents are also used in *in vivo* studies. Rats are often selected for haemodynamic tests while guinea pigs are used to study the monophasic action potential duration (MAPD) [92-94]. However, guinea pigs are unlikely to exhibit TdP because of their large repolarisation reserve, which makes this model not as sensitive as other species.

1. 4. 3 ESTABLISHING TRANSLATIONAL CORRELATIONS

Computational methods

There has been increasing awareness in the importance of translating non-clinical findings to clinical situations. Attempts are made to understand the construct validity of experimental settings and their implications for the interpretation of different study results. While it is necessary to investigate the exposure at which adverse events occur, it is not sufficient to ensure accurate prediction of the clinical outcome. In real life conditions, many other factors contribute to the manifestation of cardiovascular safety issues such as underlying comorbidities and concomitant medications. These conditions cannot be investigated extensively in a preclinical experiment or in a clinical trial. Therefore, other methods need to be established to simulate a variety of situations or relevant scenarios. One such technology is a computational model of cell electrophysiology for Purkinje fibres, in which information about the human genes responsible for individual ion currents is connected to whole cell behaviour [95]. Model simulations of selective ion channel blockade can reproduce results observed in pharmacological challenges characteristic of isolated Purkinje fibres *in vitro,* which has the potential to aid the translation between *in vitro* experiments and clinical response. More recently, Bottino developed mathematical models that use hERG IC50 data and APD results measured from dog Purkinje fibres to predict drug interaction with other cardiac ion currents and dispersion of repolarisation in transmural ECG [96]. Furthermore, this *in silico* approach allows investigation of the influence of known clinical risk factors for arrhythmia (e.g. hypokalaemia) on the pro-arrhythmic risk in patients. Mathematical models have also been used to predict the clinical torsadogenic risk based on the pre-clinical data generated early on in the development process and explore the implications of multiple-ion-channel blockade [97].

While the aforementioned experimental and computational methods represent an advancement in translational research for cardiovascular safety, protocol design and data analysis techniques can affect the quality of the data produced and can therefore ultimately influence the interpretation of cardiovascular risk [98]. The discrepancy in results due to differences in measurement methods has been demonstrated by the work done by Ollerstam et al [99] (Figure 9).

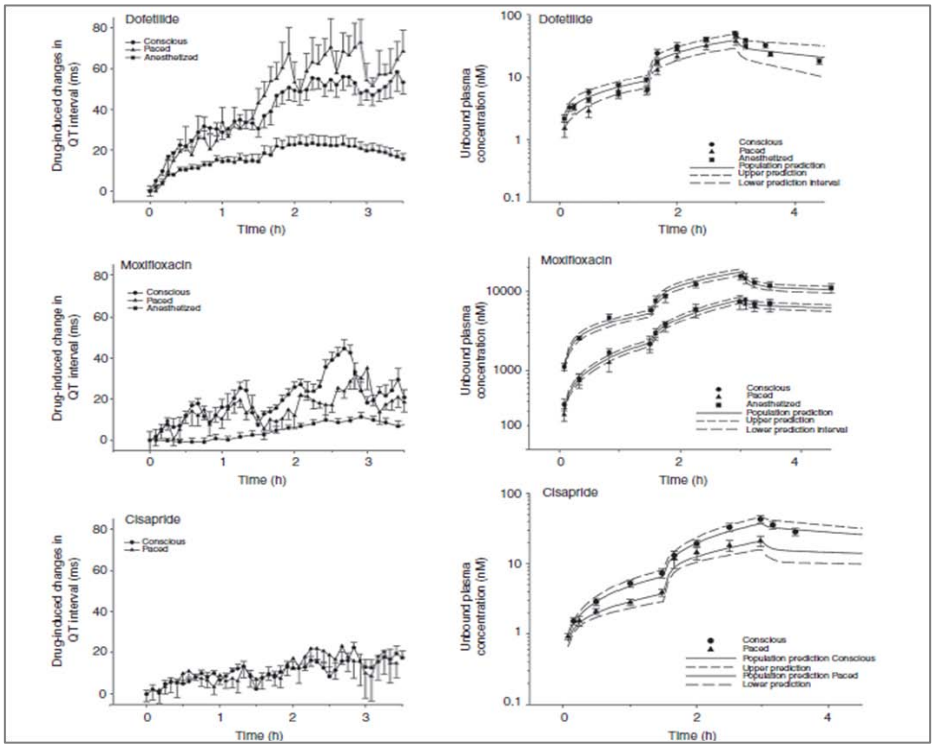


Figure 9. Left panels - the drug-induced QT-interval prolongation (mean±SE) during the initial 3.5 h of the experiments for dofetilide, moxifloxacin, and cisapride in dog models. Right panels: Plasma concentration (mean±SE) versus time for dofetilide, moxifloxacin, and cisapride in all dog models. The population pharmacokinetic prediction and the 95% prediction interval are shown superimposed on the data [95]. The various measurement methods used (in conscious, paced and anesthetised animals) render discrepancies of results.

In contrast to empirical methods, the use of a model-based approach in which *in vitro* and *in vivo* concentration-effect (PKPD) relationships are characterised, could be the foundation for translational purposes, that is, the work done by Mittelstadt and Hart (Figure 10) [100]. However, the

implementation of a model-based approach does require further consideration about the generalisability and discrepancies in PKPD relationships between species. In particular, the identification of a species showing metabolic profile comparable to humans and understanding of the differences in baseline cardiovascular parameters are crucial for predicting clinical outcome. Furthermore, having common PKPD parameters between species would facilitate the translation of results based on mechanism-based modelling techniques which allow discrimination between drug- and system-specific properties. These features reflect the inherent modularity of the method [101]. This gives rise to the significance of establishing harmonised experimental protocols so that only drug-specific parameters are evaluated while system-specific properties remain constant in the same species.

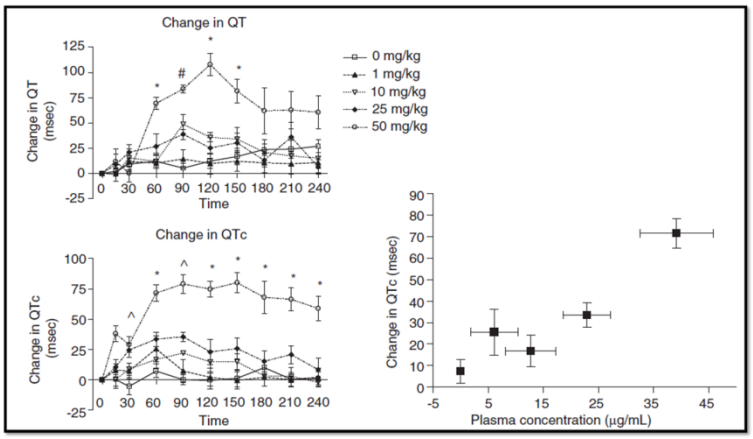


*Figure 10. An example of the use of a model-based approach in which in vitro and in vivo concentration-effect relationships are characterised and could be the foundation for translational purposes. Right panel – Change in QT and QTc from predose. Plots of mean and SEM (n = 4) QT-interval and QTc after a 60 min infusion of vehicle and four doses of moxifloxacin. *50 mg/kg significantly different (P < 0.05) than vehicle. ^25 and 50 mg/kg significantly different (P < 0.05) than vehicle. #10, 25, 50 mg/kg significantly different (P < 0.05) than vehicle. Left panel – Plasma concentration vs. change in QTc. Plot of mean and SEM (n = 4) plasma levels vs. the change in QTc-interval at the end of infusion of vehicle and four doses of moxifloxacin [96].*

Pharmacokinetic-pharmacodynamic modelling

Population pharmacokinetic (pop PK) and pharmacokinetic-pharmacodynamic (pop PKPD) models basically comprise three main components: a structural model which describes pharmacokinetics or pharmacodynamic characteristics (e.g. 2-compartment disposition model or

sigmoid E_{max} pharmacodynamic model); a statistical model describing between-subject differences and an error model which accounts for the residual variability [102]. Population models also incorporate the effect of influential covariates, such as weight, age, and pharmacogenetics. Rather than correlating the effects of these factors directly to the observed variables, data analysis relies on the assessment of covariate effect on model parameters; for example, absorption and clearance.

Empirical models have significant limitations when it comes to extrapolating PK and PD properties between species [103]. This is because the behaviour of drugs may change dramatically between different systems, be they normal or pathological. Moreover, it is increasingly being recognised that, even across *in vitro* assays, compounds may display "pluridimentional efficacy" [104]. In the past, different PKPD models were required to describe pre-clinical and clinical data [99]. In other cases, additional *in vitro* data was required to correlate findings across species [105]. Thus, the choice of parameterisation in these analysis is paramount in the success in translating results throughout drug development.

The use of a model-based approach in the evaluation of safety has several advantages, one of which is the ability to update our knowledge of NCEs [106]. Logically, one should endeavour to apply the information learnt from one study in drug development to the next. PKPD modelling has already proven to add considerable value for informed decision making regarding the safety profile of drug candidates [107]. It is a part of the pre-clinical integrated risk assessment where interspecies extrapolations are facilitated and clinical trial designs are optimised [108]. Another advantage is that the use of population methods allows for exploitation of sparse sampling of pharmacokinetic data, often collected during *in vivo* studies. This enables concurrent characterisation of the concentration *versus* time profile as well as the corresponding variability in the population [109]. Subsequently, it is possible to establish correlations between concentration-effect relationships across species. PKPD modelling can also take into consideration hysteresis and other time-dependent factors affecting the concentration-effect curve, which need to be accounted for to accurately assess the magnitude of drug-induced ECG changes [110]. Publications on interspecies correlations for cardiovascular effects are limited and the relationship of exposure and response across species is not entirely understood [108]. However, recent examples suggest that allometric scaling of data obtained in preclinical

models may be applicable to predict not only pharmacokinetics, but also drug effects in humans [101, 111]. In addition to characterising inter-species differences, the use of PKPD modelling ultimately enables the assessment of *in vitro-in vivo* correlations. In fact, clinical QT prolongation has been predicted from hERG data [105]; however, this might be limited to selective hERG-blockers.

With the introduction of mechanistic approaches and semi-mechanistic PKPD models, quantitative analysis of the dynamic interactions between drugs and biological system can be achieved [112-114]. By separating drug-specific and biological system-specific parameters, a framework can be established for translational research that links the interactions between drug, pharmacological targets and integrated disease systems in a quantitative manner [101, 115, 116]. It can be anticipated that understanding of PKPD relationships will represent a strong platform from which to explore the wider question of which experimental protocol conditions, parameter(s) and species are most suitable for translating drug effects from animals to humans [117].

The ability of the *in vitro* models to quantitatively predict drug effects at therapeutic concentrations is often limited. The current non-clinical paradigm employs a one-dimensional decision making scheme where compounds are discontinued based solely on torsadogenicity without characterisation of the clinical relevance and concentration range at which the toxicities occur. The danger is that this represents a high rate of false-positive results and perfectly efficacious compounds are not further developed. Hence, *in vitro* experimental results should be considered together with *in vivo* study results before decisions are made. The ICH S7B suggests that *in vivo* assessment of the drug effects on QT-interval duration, performed in addition to a hERG channel assay, should form the minimal basis for the evaluation of pro-arrhythmic risk. However, the utility of non-clinical assays remains questionable. This is partly due to variability in methods, species, and lack of consistency in the experimental procedures reported in the literature [85]. The results of ongoing efforts have therefore led to great improvement in the sensitivity but not the predictability of screening protocols. To establish *in vitro-in vivo* correlations, non-clinical studies must be calibrated against clinical data. Furthermore, clinically relevant endpoints must be used as a benchmark.

	Technology	Advantage	Limitation
1	*In vitro* experiments	- Able to detect the degree of QTc prolongation - Assays give an insight into the mechanisms underlying the genesis of arrhythmias - Measure action potential durations which are found to be more predictive of proarrhythmic potential drugs	- Not necessarily a good predictor to TdP - Lack of correlation between the concentration required for hERG IC50 and 10% prolongation in APD90
2	*In vivo* studies	- Dogs are sensitive in predicting even small QT prolongation - Guinea pigs possess specific ion channels similar to those in humans - Monkeys are considered to be a sensitive and useful model for assessing the potential for drug-induced QT prolongation - Mini-pigs share many of their features with humans, including anatomy, physiology and biochemistry	- Conscious animals can be constrained by the upper-limit of dose - Could have great repolarisation reserve so are not good for models of arrhythmia and not sensitive enough to detect known torsadogenic drugs - Ethical and "three Rs" issues surrounding the use of monkeys - Mini-pigs can remain in a state of excitement for several hours after feeding, which can lead to misinterpretation of CV data
3	Establishing translational correlations	- Creating models have the ability to update our knowledge of NCEs and add considerable value for informed decision making - Population methods allow for exploitation of sparse sampling of PK data - Simulations can be used to extensively test different real life scenarios from various phases of drug development - Drug- and system-specific parameters can be derived across different species	- Requires thorough understanding of pharmacology and PKPD modelling concepts - It can be challenging as routine technique due to the lack of integrated tools and methodologies

Table 1. Summary of current techniques and methodologies used for the assessment of cardiovascular safety

1.5 TRANSLATIONAL EFFORTS: CLINICAL TRIALS TO REAL-LIFE CONDITIONS

A summary of the current techniques and experimental conditions used in clinical research is presented below, with emphasis on the lack of methodology that ensures the assessment of pharmacokinetic-pharmacodynamic relationships as well as of other intrinsic and extrinsic factors known to affect heart conductivity. This is followed by 1) a discussion on the importance of characterising the potential differences between well controlled clinical trials and real life patient populations as well as 2) a summary of the main methodological aspects.

Both from a scientific and regulatory perspective, the crucial question remains is "How efficient and reliable are the pre-approval clinical trials in identifying the clinical risk of TdP, given the 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" [118]. For a variety of reasons, clinical trials may or may not detect the frequency and intensity of QT-interval prolongation but they are most unlikely to detect the risk of TdP. Firstly, the present approach to clinical trials is highly efficacy-orientated. Secondly, the number of patients exposed to the drug in these pre-approval clinical trials is not large enough to detect a relatively rare but potentially fatal risk. Thirdly, many subgroups of patients most at risk of TdP during the uncontrolled clinical use to the drug in question are usually excluded from these trials. These subgroups include: 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 the potassium channels, those susceptible to bradycardia or electrolyte imbalance, or those receiving drugs with a potential for pharmacokinetic or pharmacodynamic interactions.

1. 5. 1 THOROUGH-QT STUDY

The primary analysis of a TQT study is based on the "double-delta" methods, which is an assessment of the upper bound of the 95% one-sided confidence interval for the largest time-matched mean effect of the drug on the QTc-

interval. The result of the assessment must exclude 10-msec increase in QTc-interval if the drug is to be deemed safe, i.e. a negative study. Requiring that the largest time-matched mean difference between the drug and placebo QTc-interval be ~5 msec or less implies that the one-sided 95% confidence interval should exclude an effect of ≥ 10msec for every single measurement. Although this is thought to be the most suitable approach, it can lead to multiplicity issues and positive bias. Another limitation of the proposed double-delta method is the absence of concentration-QT relationships. From a scientific perspective, the evidence gathered to date means that any regulatory review of QT study is not complete without an assessment of concentration-QTc relationship.

1. 5. 2 PHARMACOVIGILANCE

Pharmacovigilance is the pharmacological science relating to the detection, assessment, understanding and prevention of adverse effects, particularly long term and short term side effects of medicines [119]. WHO established its Programme for International Drug Monitoring in response to the thalidomide disaster detected in 1961. Often, epidemiological concepts are used to determine the distribution and patterns of health-events, health-characteristics and their causes or influences in well-defined populations. The term epidemiology is widely applied to cover the description and causation of not only epidemic disease, but of disease in general, and even many non-disease health-related conditions, including QT/QTc-interval prolongation.

1. 5. 3 EPIDEMIOLOGICAL COHORTS AS RISK ASSESSMENT TOOLS

In 2005, the FDA published guidance for the industry on "Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment" [120]. The idea behind this stems from the fact that risk assessment during drug development on one hand should be conducted in a thorough and rigorous manner but it is also well understood that it is impossible to identify all safety concerns during clinical trials. When a compound is marketed, the number of patients is dramatically increased including those with comorbidities and on concomitant medications. Thus, it is useful and important to continue collecting post-marketing safety data and perform risk assessment based on observational data to characterise and amend the risk

profile of the product in Phase IV studies (Figure 11) . In fact, many governments and academic institutions have set up formal observational databases to monitor drug adverse events and progression of health or disease in general or specific populations for public health and research purposes. As a result, there exist several large population-based studies evaluating the association between the heart-rate corrected QT-interval and mortality [121-127].

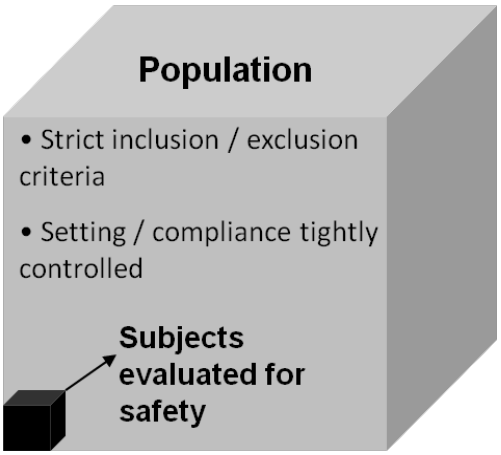


Figure 11. The black cube represent a fraction of the general population that is enrolled into clinical trials. Such population is tightly controlled by strict inclusion / exclusion criteria. Setting and compliance of trials are also strictly controlled. Due to the limitation in the trial population, it is important to continue collecting post-marketing safety data and perform risk assessment to amend the risk profile of the product in Phase IV studies.

Various types of epidemiologic studies can be used to assess if there is an association between exposure of a drug and a particular outcome [128]. In cohort studies, the investigator defines two or more groups of people that are free of a particular condition/ disease and that differ according to the extent of their exposure to a potential cause of the condition. The incidence times and rates of development are then measured and compared. Alternatively, case-control studies can be conducted to identify factors that may contribute to a medical condition / disease or treatment effect by comparing subjects who have the condition (cases) with those who do not have the condition (control) but are otherwise similar. Furthermore, medical condition / disease or treatment effect can be evaluated prospectively or retrospectively based on the timing relative to when the investigators begin to observe the development of condition / disease (Figure 12).

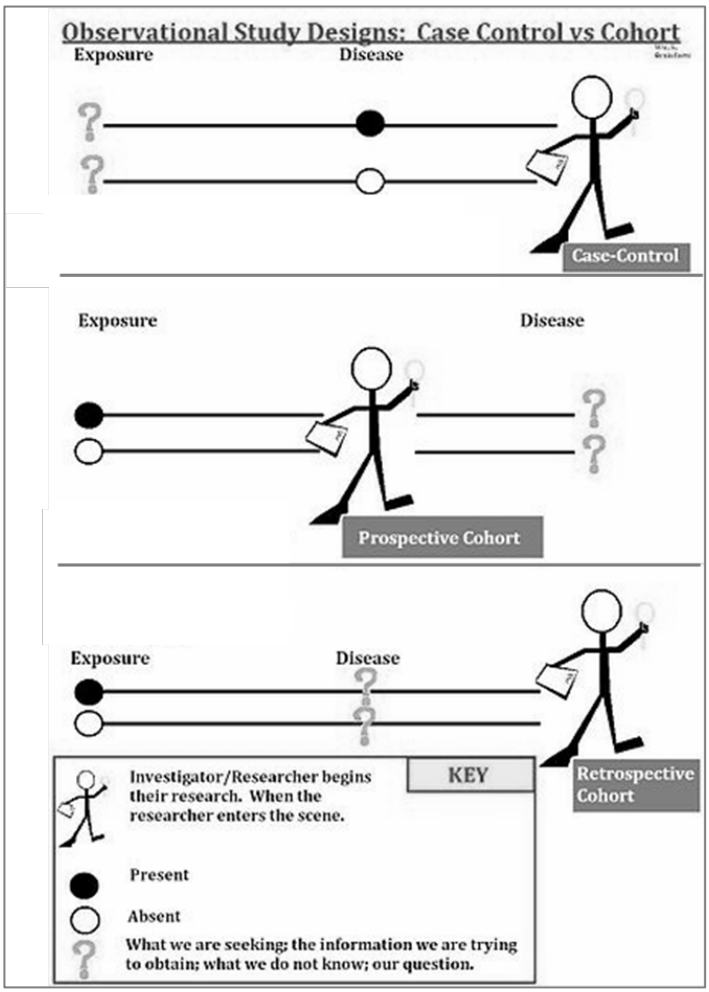


Figure 12. Case-control studies vs. Cohort studies on a timeline. http://www.wikipedia.org/wiki/Case-control_stud, as accessed on 21 April, 2012 at 9:47 EST.

Some epidemiologic studies have found results consistent with findings from the clinical trials. For example, the Ziprasidone Observational Study of Cardiac Outcomes (ZODIAC) was a post-marketing study done to measure non-suicide mortality in the year after initiation of assigned treatments. Despite known risk of QTc prolongation with ziprasidone treatment, the findings from the study failed to show that the drug is associated with an elevated risk of non-suicidal mortality in real-world use relative to olanzapine [129]. Conversely, some Phase IV studies have found inconsistencies between clinical trial outcomes and real life use. Examples of

such discrepancies include tacrine, simvastatin and celecoxib [130]. From a clinical perspective, the hiatus between clinically relevant population and clinical trial population has been known to exist, but such differences are often assumed to have negligible impact on safety and efficacy. Many studies have highlighted that many subgroups of patient cardiovascular patients are under-represented in pre-registration Phase III trials [131, 132].

1. 5. 4 ESTABLISHING TRANSLATIONAL CORRELATIONS

The aforementioned hiatus between clinical trials and epidemiological cohorts has driven a series of efforts in pharmacoepidemiology and risk management, which are aimed at the generalisation of clinical findings into real life scenarios. The assumption of generalisability is however violated when events are rare or occur at very low frequency. Specifically with regard to cardiovascular safety, the low incidence of TdP in clinical trials causes underestimation of this adverse event, which often are not identified until after the drug has been approved. This is one of the reasons to establish Phase IV studies to monitor the safety issues of new drugs. However, Phase IV studies are epidemiological in nature where the objective is to understand the risk, not necessarily to prevent or mitigate it. Most importantly, the findings are not necessarily as accurate given that exposure-response relationships are not taken into account. Moreover, epidemiology studies often use dose as a proxy for drug exposure, irrespective of evidence of differences in pharmacokinetics (e.g. bio-availability and other possible drug-drug interactions). Assumptions are also made with regard to the effects of poor compliance or non-adherence to therapy.

Bayesian hierarchical modelling

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. The key difference between classical and Bayesian statistics is the interpretation of probability. In the classical framework, probability is considered a surrogate for frequency, hence the term 'frequentist' statistics. This means that a rigid statistician can only make probability statements about repeating events, such as coin tosses, but not about events that occur only once, such as the weather tomorrow or the outcome of a particular clinical trial. This is somewhat circumvented by applying statistical methods many times, so that the frequency is in the use of

the test rather than in the repetition of the experiment. For example: a confidence interval will contain the mean in 95% of the cases in which it is constructed (and not, as is commonly believed, contain the mean with 95% probability). In a Bayesian context, probability is considered to be a quantification of 'degree of belief'. This means that Bayesian statisticians can make direct probability statements about any event, such as the probability that a treatment is superior to placebo.

The statistical reasoning behind this is that frequentist statistics relies on the likelihood, i.e. the probability of observing the data given a range of parameter values. Bayesian statistics however relies on the posterior distribution, which is the distribution of the parameters (i.e., parameter space) given the observed data. To obtain the posterior distribution, both the likelihood and a prior distribution are required. The prior distribution can include all relevant information that is not captured by the actual experiment. Since the use of this prior means that no two analyses will be the same, it is important to be explicit about the prior distribution that has been used in any Bayesian analysis. An advantage of the resulting posterior distributions is that the interpretation of the probability is more direct. Indeed, 95% credible intervals (as they are referred to in a Bayesian context) do have the interpretation that there is a 95% probability of an interval containing the true value.

The increase in the use of Bayesian statistics in recent years is explained by developments in computational sciences and enhanced computing capacity of computer processors. For most Bayesian problems a closed-form analytical solution is not possible. However, using Markov Chain Monte Carlo (MCMC) methods, it is possible to sample directly from the posterior distribution. Typically, at least 10000 samples are required to characterise the posterior distribution. Increasing computer processor speeds have made this task feasible. It is important to emphasise that in the absence of an analytical solution, MCMC does yield very good, unbiased results, as opposed to many of the maximum-likelihood based algorithms that are currently used. The development of WinBUGS [133, 134] has removed many of the hurdles associated with the development of algorithms and paved the road for Bayesian statistics in many different areas.

Another important aspect of the application of hierarchical (population) models is the focus on the population rather than the individual as the object

of the investigation. The approach is particularly suitable when information on individual subjects is limited. Regardless of the density of the data, population models describing concentration-(adverse)effect relationships should be considered an ideal tool to translate between species (i.e., non-clinical to clinical data) and between clinical trials and real life population.

Lastly, it should be noted that the availability of hierarchical models offers an important opportunity as study optimisation tool (e.g., dose selection, sampling times, treatment duration and population size). These models can also be used to support prediction and extrapolation of data across different age-groups, dosing regimens and formulations or delivery forms. Moreover, population models may enable extrapolation of long-term efficacy and safety based on short-term pharmacokinetic and treatment response data.

Clinical trial simulations

In contrast to meta-analysis, clinical trial simulation (CTS) allows for the investigation of the impact of a range of design characteristics on the power to detect a treatment effect prior to exposing patients to an experimental drug (Figure 13). In a field where most clinical trials have a conservative design, this methodology offers a unique opportunity to evaluate innovative designs. Rather than performing power calculations which only take sample size and endpoint variability into account, CTS allows calculation of power taking into account a multitude of other factors.

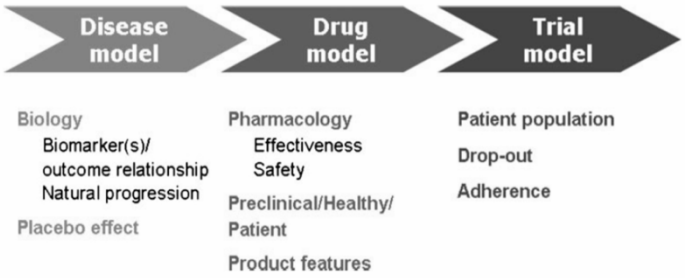


Figure 13. Major components of a clinical trial simulation (CTS). In model-based drug development, a CTS can be used to characterise the behaviour of the compound in different population groups by utilising patient-specific covariate values [135].

In general, CTS in the context of translating results from clinical trial to real life population utilises two types of models [136]. First, a drug-action

(PKPD) model is considered, which comprises pharmacokinetic and pharmacodynamic factors. In chronic diseases the model also accounts for disease progression. Unfortunately, the lack of knowledge about the mechanisms underlying treatment response in many therapeutic indications has prevented the development of mechanistic models. Hence, examples often refer to standard statistical models, such as e.g., the mixed model for repeated measures (MMRM). Such statistical models have however a downside in that they often do not incorporate concentration-effect relationships. Secondly, CTS requires a trial execution model. These models simulate other important aspects of the trial, such as dropout, compliance and protocol deviations. Therefore, one can determine all possible outcomes under candidate trial designs, allowing such trial designs to be compared in a strictly quantitative manner. Thus far, very few examples exist in which relevant design factors have been evaluated prospectively as part of the planning of a cardiovascular safety trial.

It is also important to stress that CTS allows investigation of factors that cannot be scrutinised by meta-analysis or empirical design. First, designs which have not been implemented cannot be included in a meta-analysis. Second, it is difficult to separate the influence of multiple design factors, whereas CTS allows evaluation of a single factor, one at a time. Although meta-analyses may provide valuable information about differences in patient populations and treatment response, it is unfortunate that many investigators consider overall publication review sufficient to gather evidence on the role of design factors, as often suggested in the discussion of meta-analysis results.

If simulated data is to be exchangeable with actual patient data, it is imperative that not only model parameters are unbiased, but that estimates of variability are also accurate. Often interpretation of statistical model results focuses on the predicted values of the treatment effect. This does not necessarily mean that response distributions reflect what occurs in the true patient population. In fact, it is not infrequent to see model misspecifications being corrected by inflated estimates of variability. It is therefore critical for clinicians to understand that standard goodness-of-fit criteria do not take simulation characteristics into account and may therefore not be indicative of the best model. Such a comparison between simulated and original data can be performed using graphical and statistical tools.

CTS also relies on the availability of accurate model parameter and corresponding distributions to investigate "what if" scenarios across a different range of conditions or design features, such as population size, stratification levels, dose range, sampling scheme, and even different endpoints. One of the main advantages of such a virtual or statistical experiment is the possibility to predict 'trial performance' and so to identify potential limitations in a study and protocol design prior to its implementation and decide if the trial should be performed. In fact, some clinical trial simulations have been evaluated against outcomes from real trials.

1.6 SUMMARY

In this chapter, we have provided an overview of the concepts and underlying physiological and methodological aspects pertinent to the evaluation of chronotropic effects on cardiac conductivity and more specifically on QT-interval. Prolonged QT/QTc-intervals can degrade into ventricular arrhythmias and even sudden death. Due to its presence prior to incidences of fatal arrhythmias, QT/QTc-interval prolongation has been chosen as a safety biomarker. Currently, a multitude of safety trials are done to test the propensity of drug-induced prolongation in different phases of the drug development process.

Despite the numerous efforts to improve the detection of QT/QTc-prolonging effects, many limitations still exist in the current approaches. As described in the aforementioned paragraphs, a major hurdle remains the lack of an integrated approach in which exposure-response relationships are used as basis for translational purposes throughout the development and life cycle of a product. Moreover, past attempts to introduce PKPD relationships have ignored the need to discriminate between drug- and system-specific parameter, making it challenging to assess risk in a quantitative manner. This limitation is accompanied by the discrepancy between the target patient population and healthy subjects currently used for the evaluation of drug-induced effects. The assumption of comparable drug response across population does not necessarily hold true. Tools are needed that enable the characterisation of drug properties, supporting the decision to progress or terminate the development a compound.

REFERENCES

[1] Lasser KE, Allen PD, Woolhandler SJ, Himmelstein DU, Wolfe SM, Bor DH. Timing of new black box warnings and withdrawals for prescription medications. JAMA 2002 May 1;287(17):2215-20.

[2] Munos B. Lessons from 60 years of pharmaceutical innovation. Nat Rev Drug Discov 2009 Dec;8(12):959-68.

[3] Valentin JP, Hammond T. Safety and secondary pharmacology: successes, threats, challenges and opportunities. J Pharmacol Toxicol Methods 2008 Sep-Oct;58(2):77-87.

[4] Laverty HG, Benson B, Cartwright EJ, Cross MJ, Garland C, Hammond TG, et al. How can we improve our understanding of cardiovascular safety liabilities to develop safer medicines? Br J Pharmacol 2011;163:675-93.

[5] Pettit SD, Berridge B, Sarazan RD. A call for more integrated cardiovascular safety assessment. J Pharmacol Toxicol Methods 2010 Jan-Feb;61(1):1-2.

[6] Krejsa CM, Horvath D, Rogalski SL, Penzotti JE, Mao B, Barbosa F, et al. Predicting ADME properties and side effects: the BioPrint approach. Curr Opin Drug Discov Devel 2003 Jul;6(4):470-80.

[7] Budnitz DS, Pollock DA, Weidenbach KN, Mendelsohn AB, Schroeder TJ, Annest JL. National surveillance of emergency department visits for outpatient adverse drug events. JAMA 2006 Oct 18;296(15):1858-66.

[8] Sagall EL, Horn C.O., Riseman, J.E.F. Studies on the action of quinidine in man: I. Measurement of the speed and duration of the effect following oral and intramuscular administration. Arch Intern Med 1943(71):460-73.

[9] Sturnick M, Riseman, J.E.F., Sagall, E.L. Studies on the action of quinidine in man: II. Intracuscular administration of a soluble preparation of quinidine in the treatment of acute cardiac arrhythmias. JAMA 1943(121):917-20.

[10] Lewis T, Drury, A.N. Revised views of refractory period in relation to drugs reputed prolong it and in relation to curcus movement. Heart 1926(13):95-100.

[11] von Frey W. Weitere erfahrungen mit chinidin bei absoluter herzunregelmassigkeit. Berliner Klin Wochenschr 1918(55):849-53.

[12] Levy RL. Clinical studies with quinidine: the clinical toxicology of quinine. JAMA 1922(79):1108-13.

[13] Gordon B, Matton, M., Levine, S.A. The mechanism of death from quinidine and a method of resuscitation: an experimental study. J Clin Invest 1925(1):497-517.

[14] Burda CD. Electrocardiographic abnormalities induced by thioridazine (Mellaril). Am Heart J 1968 Aug;76(2):153-6.

[15] Desautels S, Filteau C, St-Jean A. Ventricular Tachycardia Associated with Administration of Thioridazine Hydrochloride (Mellaril): Report of a Case with a Favourable Outcome. Can Med Assoc J 1964 Apr 25;90:1030-1.

[16] Giles TD, Modlin RK. Death associated with ventricular arrhythmia and thioridazine hydrochloride. JAMA 1968 Jul 8;205(2):108-10.

[17] Graupner KI, Murphree OD, Meduna LJ. Electrocardiographic Changes Associated with the Use of Thioridazine. J Neuropsychiatr 1964 Aug;5:344-50.

[18] Hollander PB, Cain RM. Effects of thioridazine on transmembrane potential and contractile characteristics of guinea pig hearts. Eur J Pharmacol 1971 Oct;16(2):129-35.

[19] Wendkos MH. The Significance of Electrocardiographic Changes Produced by Thioridazine. J New Drugs 1964 Nov-Dec;40:322-32.

[20] Wendkos MH, Thornton CG. An electrocardiographic survey of thioridazine treated patients. Behav Neuropsychiatry 1969 Dec;1(9):18-22.

[21] Ban TA, St Jean A. Electrocardiographic changes induced by phenothiazine drugs. Am Heart J 1965 Oct;70(4):575-6.

[22] Ban TA, Stjean A. The Effect of Phenothiazines on the Electrocardiogram. Can Med Assoc J 1964 Sep 5;91:537-40.

[23] Hollister LE, Kosek JC. Sudden Death during Treatment with Phenothiazine Derivatives. JAMA 1965 Jun 21;192:1035-8.

[24] Huston JR. Electrocardiographic changes produced by thioridazine and chlorpromazine. Am Heart J 1969 May;77(5):713-4.

[25] Lingjaerde O. Electrocardiographic changes, disturbances of cardiac rhythm, and sudden deaths during treatment with phenothiazine drugs. Tidsskr Nor Laegeforen 1967 Jan 15;87(2):90-4.

[26] Richardson HL, Graupner KI, Richardson ME. Intramyocardial lesions in patients dying suddenly and unexpectedly. JAMA 1966 Jan 24;195(4):254-60.

[27] Wendkos MH. Cardiac changes related to phenothiazine therapy, with special reference to thioridazine. J Am Geriatr Soc 1967 Jan;15(1):20-8.

[28] Towbin JA, Wang Z, Li H. Genotype and severity of long QT syndrome. Arch Pathol Lab Med 2001 Jan;125(1):116-21.

[29] Berul CI. Congenital long-QT syndromes: who's at risk for sudden cardiac death? Circulation 2008 Apr 29;117(17):2178-80.

[30] Moss AJ, Schwartz PJ, Crampton RS, Tzivoni D, Locati EH, MacCluer J, et al. The long QT syndrome. Prospective longitudinal study of 328 families. Circulation 1991 Sep;84(3):1136-44.

[31] Schwartz PJ, Zaza A, Locati E, Moss AJ. Stress and sudden death. The case of the long QT syndrome. Circulation 1991 Apr;83(4 Suppl):II71-80.

[32] Bazett HC. (1920). "An analysis of the time-relations of electrocardiograms". Heart (7): 353–370.

[33] Fridericia LS (1920). "The duration of systole in the electrocardiogram of normal subjects and of patients with heart disease". Acta Medica Scandinavica (53): 469 –486.

[34] Dessertenne F. [Ventricular tachycardia with 2 variable opposing foci]. Arch Mal Coeur Vaiss 1966 Feb;59(2):263-72.

[35] Points to consider: the assessment of the potential for QT interval prolongation by non-cardiovascular medicinal products. In: EMEA, editor. London, UK; 1997.

[36] Morganroth J. Relations of QTc prolongation on the electrocardiogram to torsades de pointes: definitions and mechanisms. Am J Cardiol 1993 Aug 26;72(6):10B-3B.

[37] Passman R, Kadish A. Polymorphic ventricular tachycardia, long Q-T syndrome, and torsades de pointes. Med Clin North Am 2001 Mar;85(2):321-41.

[38] Stratmann HG, Kennedy HL. Torsades de pointes associated with drugs and toxins: recognition and management. Am Heart J 1987 Jun;113(6):1470-82.

[39] Tzivoni D, Keren A, Banai S, Stern S. Terminology of torsades de pointes. Cardiovasc Drugs Ther 1991 Apr;5(2):505-7.

[40] Salle P, Rey JL, Bernasconi P, Quiret JC, Lombaert M. [Torsades de pointe. Apropos of 60 cases]. Ann Cardiol Angeiol (Paris) 1985 Jun;34(6):381-8.

[41] Milon D, Daubert JC, Saint-Marc C, Gouffault J. [Torsade de pointes. Apropos of 54 cases]. Ann Fr Anesth Reanim 1982;1(5):513-20.

[42] Fung MC, Hsiao-hui Wu, H., Kwong, K., Hornbuckle, K., Muniz, E. Evaluation of the profile of patients with QTc prolongation in spontaneous adverse event reporting over the past three decades - 1969-98. Pharmacoepidemiol Drug Safety 2000;9(Suppl. 1):S24.

[43] Shah RR. Withdrawal of terodiline: A tale of two toxicities. London: Wiley & Sons Ltd; 2001.

[44] Roden DM. Drug-induced prolongation of the QT interval. N Engl J Med 2004 Mar 4;350(10):1013-22.

[45] Yang T, Snyders D, Roden DM. Drug block of I(kr): model systems and relevance to human arrhythmias. J Cardiovasc Pharmacol 2001 Nov;38(5):737-44.

[46] Bauman JL, Bauernfeind RA, Hoff JV, Strasberg B, Swiryn S, Rosen KM. Torsade de pointes due to quinidine: observations in 31 patients. Am Heart J 1984 Mar;107(3):425-30.

[47] MacNeil DJ, Davies RO, Deitchman D. Clinical safety profile of sotalol in the treatment of arrhythmias. Am J Cardiol 1993 Aug 12;72(4):44A-50A.

[48] Throckmorton, D. Cardiovascular Drugs That Prolong The QT Interval, Division of Cardio-Renal Drug Products, FDA , July 2000.

[49] The clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs: preliminary concept paper. 2002.

[50] Guidance for industry: ICH E14 The clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs. U.S. Department of Health and Human Services. Food and Drug Administration. (2005)

[51] Pollard CE, Abi Gerges N, Bridgland-Taylor MH, Easter A, Hammond TG, Valentin JP. An introduction to QT interval prolongation and non-clinical approaches to assessing and reducing risk. Br J Pharmacol 2010 Jan;159(1):12-21.

[52] Ducroq J, Printemps R, Guilbot S, Gardette J, Salvetat C, Le Grand M. Action potential experiments complete hERG assay and QT-interval measurements in cardiac preclinical studies. J Pharmacol Toxicol Methods 2007 Sep-Oct;56(2):159-70.

[53] Vargas HM, Bass AS, Breidenbach A, Feldman HS, Gintant GA, Harmer AR, et al. Scientific review and recommendations on preclinical cardiovascular safety evaluation of biologics. J Pharmacol Toxicol Methods 2008 Sep-Oct;58(2):72-6.

[54] Milberg P, Ramtin S, Monnig G, Osada N, Wasmer K, Breithardt G, et al. Comparison of the in vitro electrophysiologic and proarrhythmic effects of amiodarone and sotalol in a rabbit model of acute atrioventricular block. J Cardiovasc Pharmacol 2004 Sep;44(3):278-86.

[55] Redfern WS, Carlsson L, Davis AS, Lynch WG, MacKenzie I, Palethorpe S, et al. Relationships between preclinical cardiac electrophysiology, clinical QT interval prolongation and torsade de pointes for a broad range of drugs: evidence for a provisional safety margin in drug development. Cardiovasc Res 2003 Apr 1;58(1):32-45.

[56] Shah RR. Drugs, QT interval prolongation and ICH E14: the need to get it right. Drug Saf 2005;28(2):115-25.

[57] Omata T, Kasai C, Hashimoto M, Hombo T, Yamamoto K. QT PRODACT: comparison of non-clinical studies for drug-induced delay in ventricular repolarization and their role in safety evaluation in humans. J Pharmacol Sci 2005;99(5):531-41.

[58] Webster R, Leishman D, Walker D. Towards a drug concentration effect relationship for QT prolongation and torsades de pointes. Curr Opin Drug Discov Devel 2002 Jan;5(1):116-26.

[59] Lawrence CL, Pollard CE, Hammond TG, Valentin JP. Nonclinical proarrhythmia models: predicting Torsades de Pointes. J Pharmacol Toxicol Methods 2005 Jul-Aug;52(1):46-59.

[60] Lee N, Authier S, Pugsley MK, Curtis MJ. The continuing evolution of torsades de pointes liability testing methods: is there an end in sight? Toxicol Appl Pharmacol 2010 Mar 1;243(2):146-53.

[61] Guidance for industry: ICH Topic S7B The nonclinical Evaluation of the Potential for delayed Ventricular Repolarization (QT Interval Prolongation) by Human Pharmaceuticals. U.S. Department of Health and Human Services. Food and Drug Administration. (2005.

[62] Holbrook M, Malik M, Shah RR, Valentin JP. Drug induced shortening of the QT/QTc interval: an emerging safety issue warranting further modelling and evaluation in drug research and development? J Pharmacol Toxicol Methods 2009 Jan-Feb;59(1):21-8.

[63] Gintant GA, Limberis JT, McDermott JS, Wegner CD, Cox BF. The canine Purkinje fiber: an in vitro model system for acquired long QT syndrome and drug-induced arrhythmogenesis. J Cardiovasc Pharmacol 2001 May;37(5):607-18.

[64] Zipes DP. Mechanisms of clinical arrhythmias. J Cardiovasc Electrophysiol 2003 Aug;14(8):902-12.

[65] Eckardt L, Haverkamp W, Borggrefe M, Breithardt G. Experimental models of torsade de pointes. Cardiovasc Res 1998 Jul;39(1):178-93.

[66] Hammond TG, Carlsson L, Davis AS, Lynch WG, MacKenzie I, Redfern WS, et al. Methods of collecting and evaluating non-clinical cardiac electrophysiology data in the pharmaceutical industry: results of an international survey. Cardiovasc Res 2001 Mar;49(4):741-50.

[67] Guth BD, Germeyer S, Kolb W, Markert M. Developing a strategy for the nonclinical assessment of proarrhythmic risk of pharmaceuticals due to prolonged ventricular repolarization. J Pharmacol Toxicol Methods 2004 May-Jun;49(3):159-69.

[68] Lu HR, Vlaminckx E, Hermans AN, Rohrbacher J, Van Ammel K, Towart R, et al. Predicting drug-induced changes in QT interval and arrhythmias: QT-shortening drugs point to gaps in the ICHS7B Guidelines. Br J Pharmacol 2008 Aug;154(7):1427-38.

[69] Martin RL, McDermott JS, Salmen HJ, Palmatier J, Cox BF, Gintant GA. The utility of hERG and repolarization assays in evaluating delayed cardiac repolarization: influence of multi-channel block. J Cardiovasc Pharmacol 2004 Mar;43(3):369-79.

[70] Terrar DA, Wilson CM, Graham SG, Bryant SM, Heath BM. Comparison of guinea-pig ventricular myocytes and dog Purkinje fibres for in vitro assessment of drug-induced delayed repolarization. J Pharmacol Toxicol Methods 2007 Sep-Oct;56(2):171-85.

[71] Milberg P, Hilker E, Ramtin S, Cakir Y, Stypmann J, Engelen MA, et al. Proarrhythmia as a class effect of quinolones: increased dispersion of repolarization and triangulation of action potential predict torsades de pointes. J Cardiovasc Electrophysiol 2007 Jun;18(6):647-54.

[72] Thomsen MB, Verduyn SC, Stengl M, Beekman JD, de Pater G, van Opstal J, et al. Increased short-term variability of repolarization predicts d-sotalol-induced torsades de pointes in dogs. Circulation 2004 Oct 19;110(16):2453-9.

[73] Antzelevitch C, Shimizu W. Cellular mechanisms underlying the long QT syndrome. Curr Opin Cardiol 2002 Jan;17(1):43-51.

[74] Hondeghem LM, Carlsson L, Duker G. Instability and triangulation of the action potential predict serious proarrhythmia, but action potential duration prolongation is antiarrhythmic. Circulation 2001 Apr 17;103(15):2004-13.

[75] Hondeghem LM, Dujardin K, De Clerck F. Phase 2 prolongation, in the absence of instability and triangulation, antagonizes class III proarrhythmia. Cardiovasc Res 2001 May;50(2):345-53.

[76] Suter W. Predictive value of in vitro safety studies. Curr Opin Chem Biol 2006 Aug;10(4):362-6.

[77] Pinney SP, Koller BS, Franz MR, Woosley RL. Terfenadine increases the QT interval in isolated guinea pig heart. J Cardiovasc Pharmacol 1995 Jan;25(1):30-4.

[78] Szilagyi S, Pollesello P, Levijoki J, Kaheinen P, Haikala H, Edes I, et al. The effects of levosimendan and OR-1896 on isolated hearts, myocyte-sized preparations and phosphodiesterase enzymes of the guinea pig. Eur J Pharmacol 2004 Feb 13;486(1):67-74.

[79] Wu L, Shryock JC, Song Y, Li Y, Antzelevitch C, Belardinelli L. Antiarrhythmic effects of ranolazine in a guinea pig in vitro model of long-QT syndrome. J Pharmacol Exp Ther 2004 Aug;310(2):599-605.

[80] Chen X, Cordes JS, Bradley JA, Sun Z, Zhou J. Use of arterially perfused rabbit ventricular wedge in predicting arrhythmogenic potentials of drugs. J Pharmacol Toxicol Methods 2006 Nov-Dec;54(3):261-72.

[81] Heath BM, Cui Y, Worton S, Lawton B, Ward G, Ballini E, et al. Translation of flecainide- and mexiletine-induced cardiac sodium channel inhibition and ventricular conduction slowing from nonclinical models to clinical. J Pharmacol Toxicol Methods May-Jun;63(3):258-68.

[82] Guidance for industry: ICH Topic S7A Safety Pharmacolog Studies for Human Pharmaceuticals. U.S. Department of Health and Human Services. Food and Drug Administration. (2000)

[83] De Clerck F, Van de Water A, D'Aubioul J, Lu HR, van Rossem K, Hermans A, et al. In vivo measurement of QT prolongation, dispersion and arrhythmogenesis: application to the preclinical cardiovascular safety pharmacology of a new chemical entity. Fundam Clin Pharmacol 2002 Apr;16(2):125-40.

[84] Hauser DS, Stade M, Schmidt A, Hanauer G. Cardiovascular parameters in anaesthetized guinea pigs: a safety pharmacology screening model. J Pharmacol Toxicol Methods 2005 Jul-Aug;52(1):106-14.

[85] Hanson LA, Bass AS, Gintant G, Mittelstadt S, Rampe D, Thomas K. ILSI-HESI cardiovascular safety subcommittee initiative: evaluation of three non-clinical models of QT prolongation. J Pharmacol Toxicol Methods 2006 Sep-Oct;54(2):116-29.

[86] Petit SD, et al. A call for more integrated cardiovascular safety assessment. J Pharmacol Toxicol Methods 2010 Jan-Feb;61(1):1-2

[87] Authier S, Tanguay JF, Fournier S, Gauvin D, Legaspi M, Chaurand F, et al. Conscious and anesthetized non-human primate safety pharmacology models: hemodynamic sensitivity comparison. J Pharmacol Toxicol Methods 2008 Sep-Oct;58(2):94-8.

[88] Ando K, Hombo T, Kanno A, Ikeda H, Imaizumi M, Shimizu N, et al. QT PRODACT: in vivo QT assay with a conscious monkey for assessment of the potential for drug-induced QT interval prolongation. J Pharmacol Sci 2005;99(5):487-500.

[89] Watson KJ, Gorczyca WP, Umland J, Zhang Y, Chen X, Sun SZ, et al. Pharmacokinetic-pharmacodynamic modelling of the effect of Moxifloxacin on QTc prolongation in telemetered cynomolgus monkeys. J Pharmacol Toxicol Methods May-Jun;63(3):304-13.

[90] Bode G, Clausing P, Gervais F, Loegsted J, Luft J, Nogues V, et al. The utility of the minipig as an animal model in regulatory toxicology. J Pharmacol Toxicol Methods Nov-Dec;62(3):196-220.

[91] Markert M, Stubhan M, Mayer K, Trautmann T, Klumpp A, Schuler-Metz A, et al. Validation of the normal, freely moving Gottingen minipig for pharmacological safety testing. J Pharmacol Toxicol Methods 2009 Sep-Oct;60(2):79-87.

[92] Fossa AA, Wisialowski T, Wolfgang E, Wang E, Avery M, Raunig DL, et al. Differential effect of hERG blocking agents on cardiac electrical alternans in the guinea pig. Eur J Pharmacol 2004 Feb 20;486(2):209-21.

[93] Hamlin RL, Kijtawornrat A, Keene BW, Hamlin DM. QT and RR intervals in conscious and anesthetized guinea pigs with highly varying RR intervals and given QTc-lengthening test articles. Toxicol Sci 2003 Dec;76(2):437-42.

[94] Testai L, Calderone V, Salvadori A, Breschi MC, Nieri P, Martinotti E. QT prolongation in anaesthetized guinea-pigs: an experimental approach for preliminary screening of torsadogenicity of drugs and drug candidates. J Appl Toxicol 2004 May-Jun;24(3):217-22.

[95] Sampson KJ, Iyer V, Marks AR, Kass RS. A computational model of Purkinje fibre single cell electrophysiology: implications for the long QT syndrome. J Physiol 2010 Jul 15;588(Pt 14):2643-55.

[96] Bottino D, Penland RC, stamps A, Traebert M, Dumotier B, Georgiva A, et al. Preclinical cardiac safety assessment of pharmaceutical compounds using an integrated systems-based computer model of the heart. Prog Biophys Mol Biol 2006 Jan-Apr;90(1-3):414-43.

[97] Mirams G.R. CY, Sher A., Noble D., et al. Simulation of multiple ion channel block provides improved early prediction of compounds'clincial torsadogenic risk. Cariovascular Research 2011.

[98] Sivarajah A, Collins S, Sutton MR, Regan N, West H, Holbrook M, et al. Cardiovascular safety assessments in the conscious telemetered dog: utilisation of super-intervals to enhance statistical power. J Pharmacol Toxicol Methods 2010 Jul-Aug;62(1):12-9.

[99] Ollerstam A, Visser SA, Duker G, Forsberg T, Persson AH, Nilsson LB, et al. Comparison of the QT interval response during sinus and paced rhythm in conscious and anesthetized beagle dogs. J Pharmacol Toxicol Methods 2007 Sep-Oct;56(2):131-44.

[100] Mittelstadt SW, Hart SM. Effects of moxifloxacin on QT interval in conscious dogs. J Vet Pharmacol Ther 2005 Jun;28(3):253-6.

[101] Danhof M, de Lange EC, Della Pasqua OE, Ploeger BA, Voskuyl RA. Mechanism-based pharmacokinetic-pharmacodynamic (PK-PD) modeling in translational drug research. Trends Pharmacol Sci 2008 Apr;29(4):186-91.

[102] Sheiner LB, Steimer JL. Pharmacokinetic/pharmacodynamic modeling in drug development. Annu Rev Pharmacol Toxicol 2000;40:67-95.

[103] Van Der Graaf PH, Gabrielsson J. Pharmacokinetic-pharmacodynamic reasoning in drug discovery and early development. Future Med Chem 2009 Nov;1(8):1371-4.

[104] Kenakin TP. Cellular assays as portals to seven-transmembrane receptor-based drug discovery. Nat Rev Drug Discov 2009 Aug;8(8):617-26.

[105] Jonker DM, Kenna LA, Leishman D, Wallis R, Milligan PA, Jonsson EN. A pharmacokinetic-pharmacodynamic model for the quantitative prediction of dofetilide clinical QT prolongation from human ether-a-go-go-related gene current inhibition data. Clin Pharmacol Ther 2005 Jun;77(6):572-82.

[106] Lalonde RL, Kowalski KG, Hutmacher MM, Ewy W, Nichols DJ, Milligan PA, et al. Model-based drug development. Clin Pharmacol Ther 2007 Jul;82(1):21-32.

[107] Cavero I. Using pharmacokinetic/pharmacodynamic modelling in safety pharmacology to better define safety margins: a regional workshop of the Safety Pharmacology Society. Expert Opin Drug Saf 2007 Jul;6(4):465-71.

[108] Langdon G, Davis JD, McFadyen LM, Dewhurst M, Brunton NS, Rawal JK, et al. Translational pharmacokinetic-pharmacodynamic modelling; application to cardiovascular safety data for PF-00821385, a novel HIV agent. Br J Clin Pharmacol 2010 Apr;69(4):336-45.

[109] Fleury A, Lave T, Jonsson F, Schmitt M, Hirkaler G, Polonchuk L, et al. A pharmacokinetic-pharmacodynamic model for cardiovascular safety assessment of R1551. J Pharmacol Toxicol Methods 2010 Jan-Feb;63(1):123-33.

[110] Ollerstam A, Visser SA, Persson AH, Eklund G, Nilsson LB, Forsberg T, et al. Pharmacokinetic-pharmacodynamic modeling of drug-induced effect on the QT interval in conscious telemetered dogs. J Pharmacol Toxicol Methods 2006 Mar-Apr;53(2):174-83.

[111] Zuideveld KP, Van der Graaf PH, Peletier LA, Danhof M. Allometric scaling of pharmacodynamic responses: application to 5-Ht1A receptor mediated responses from rat to man. Pharm Res 2007 Nov;24(11):2031-9.

[112] Agoram BM, Martin SW, van der Graaf PH. The role of mechanism-based pharmacokinetic-pharmacodynamic (PK-PD) modelling in translational research of biologics. Drug Discov Today 2007 Dec;12(23-24):1018-24.

[113] Gobburu JV, Lesko LJ. Quantitative disease, drug, and trial models. Annu Rev Pharmacol Toxicol 2009;49:291-301.

[114] Danhof M, Van der Graaf, P.H., Jonker, D.M., Visser, S.A.G., Zuiderveld, K.P. Mechanism-based pharmacokinetic-pharmacodynamic concentration-effect relationhips - application in drug candidate selection and lead optimization. In: Testa B, Van de Waterbeemd, H, editor. Comprehensive Medicinal Chemistry II (Part 5). UK: Elsevier; 2007. p. 885-908.

[115] Mager DE, Jusko WJ. Development of translational pharmacokinetic-pharmacodynamic models. Clin Pharmacol Ther 2008 Jun;83(6):909-12.

[116] Ploeger BA, van der Graaf PH, Danhof M. Incorporating receptor theory in mechanism-based pharmacokinetic-pharmacodynamic (PK-PD) modeling. Drug Metab Pharmacokinet 2009;24(1):3-15.

[117] Krishna R. Model-based evaluation of QTc interval risk: an increasing emphasis on early decision making. J Clin Pharmacol 2009 Sep;49(9):1010-1.

[118] Bonate PL, Russell T. Assessment of QTc prolongation for non-cardiac-related drugs from a drug development perspective. J Clin Pharmacol 1999 Apr;39(4):349-58.

[119] WHO. The Importance of Pharmacovigilance. United Kingdom; 2002.

[120] Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment. In: Center for Drug Evaluation and Research (CDER), editor; 2005.

[121] de Bruyne MC, Hoes AW, Kors JA, Hofman A, van Bemmel JH, Grobbee DE. Prolonged QT interval predicts cardiac and all-cause mortality in the elderly. The Rotterdam Study. Eur Heart J 1999 Feb;20(4):278-84.

[122] Dekker JM, Schouten EG, Klootwijk P, Pool J, Kromhout D. Association between QT interval and coronary heart disease in middle-aged and elderly men. The Zutphen Study. Circulation 1994 Aug;90(2):779-85.

[123] Elming H, Holm E, Jun L, Torp-Pedersen C, Kober L, Kircshoff M, et al. The prognostic value of the QT interval and QT interval dispersion in all-cause and cardiac mortality and morbidity in a population of Danish citizens. Eur Heart J 1998 Sep;19(9):1391-400.

[124] Goldberg RJ, Bengtson J, Chen ZY, Anderson KM, Locati E, Levy D. Duration of the QT interval and total and cardiovascular mortality in healthy persons (The Framingham Heart Study experience). Am J Cardiol 1991 Jan 1;67(1):55-8.

[125] Montanez A, Ruskin JN, Hebert PR, Lamas GA, Hennekens CH. Prolonged QTc interval and risks of total and cardiovascular mortality and sudden death in the general population: a review and qualitative overview of the prospective cohort studies. Arch Intern Med 2004 May 10;164(9):943-8.

[126] Okin PM, Devereux RB, Howard BV, Fabsitz RR, Lee ET, Welty TK. Assessment of QT interval and QT dispersion for prediction of all-cause and cardiovascular mortality in American Indians: The Strong Heart Study. Circulation 2000 Jan 4-11;101(1):61-6.

[127] Robbins J, Nelson JC, Rautaharju PM, Gottdiener JS. The association between the length of the QT interval and mortality in the Cardiovascular Health Study. Am J Med 2003 Dec 15;115(9):689-94.

[128] Rothman K.J. G, S., Lash, T.L. Types of Epidemiologic Studies. Modern Epidemiology, 3rd Edition: Lippincott Williams & Wilkins; 2008.

[129] Strom BL, Eng SM, Faich G, Reynolds RF, D'Agostino RB, Ruskin J, et al. Comparative mortality associated with ziprasidone and olanzapine in real-world use among 18,154 patients with schizophrenia: The Ziprasidone Observational Study of Cardiac Outcomes (ZODIAC). Am J Psychiatry 2011 Feb;168(2):193-201.

[130] Martin K, Bégaud, B., Latry, P., Miremont-Salamé, G., Fourrier, A., Moore, N. Differences between clinical trials and postmarketing use. Br J Clin Pharmacol 2003;57(1):86-92.

[131] Verdoux H, Begaud B. Pharmaco-epidemiology: what do (and don't) we know about utilisation and impact of psychotropic medications in real-life conditions? Br J Psychiatry 2004 Aug;185:93-4.

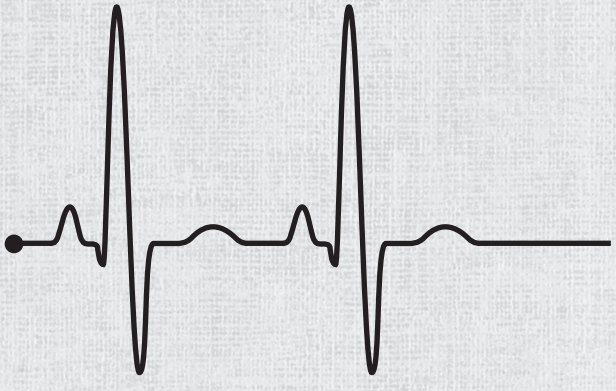
[132] Wieringa NF, Vos R, van der Werf GT, van der Veen WJ, de Graeff PA. Co-morbidity of 'clinical trial' versus 'real-world' patients using cardiovascular drugs. Pharmacoepidemiol Drug Saf 2000 Dec;9(7):569-79.

[133] Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS - A Bayesian modelling framework: concepts, structure and extensibility. Stat Comput 2000(10):325-37.

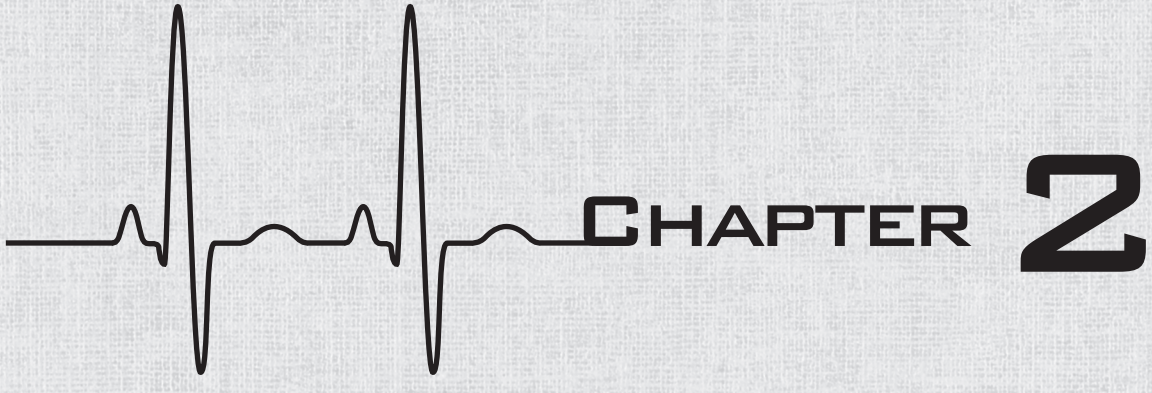
[134] Lunn DJ, Best N, Thomas A, Wakefield J, Spiegelhalter D. Bayesian analysis of population PK/PD models: general concepts and software. J Pharmacokinet Pharmacodyn 2002 Jun;29(3):271-307.

[135] Bellanti F, Della Pasqua O. Modelling and simulation as research tools in paediatric drug development. Eur J Clin Pharmacol 2011 May;67 Suppl 1:75-76.

[136] Holford NH, Kimko HC, Monteleone JP, Peck CC. Simulation of clinical trials. Annu Rev Pharmacol Toxicol 2000;40:209-34.



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-arrhythmic 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 dromotropic 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_{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 pro-active 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 $\geq$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 $\geq$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 supra-therapeutic 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

[1] Pater C. Methodological considerations in the design of trials for safety assessment of new drugs and chemical entities. Curr Control Trials Cardiovasc Med 2005 Feb 3;6(1):1.

[2] Guidance for Industry: ICH E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrythmic Potential for Non-Antiarrythmic Drugs. U.S. Department of Health and Human Services. Food and Drug Administration. (2005)

[3] Danhof M, de Lange EC, Della Pasqua OE, Ploeger BA, Voskuyl RA. Mechanism-based pharmacokinetic-pharmacodynamic (PK-PD) modeling in translational drug research. Trends Pharmacol Sci 2008 Apr;29(4):186-91.

[4] Piccini JP, Whellan DJ, Berridge BR, Finkle JK, Pettit SD, Stockbridge N, et al. Current challenges in the evaluation of cardiac safety during drug development: translational medicine meets the Critical Path Initiative. Am Heart J 2009 Sep;158(3):317-26.

[5] Turner JR. Drug safety, medication safety, patient safety: An overview of recent FDA guidances and initiatives. Regulatory Rapporteur 2009;6(4).

[6] Kowey PR, Malik, M. The QT interval as it relates to the safety of non-cardiac drugs. European Heart Journal Supplements 2007;9(Supplement G).

[7] Warn DE, Thompson SG, Spiegelhalter DJ. Bayesian random effects meta-analysis of trials with binary outcomes: methods for the absolute risk difference and relative risk scales. Stat Med 2002 Jun 15;21(11):1601-23.

[8] Bellanti F, Della Pasqua O. Modelling and simulation as research tools in paediatric drug development. Eur J Clin Pharmacol 2011 May;67 Suppl 1:75-86.

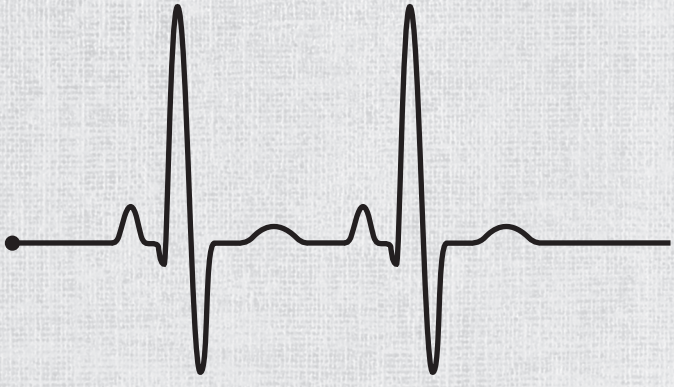
[9] Vickers A. A basic introduction to medical research. Part II: An overview of different research methods. Complement Ther Nurs Midwifery 1995 Aug;1(4):113-7.

[10] Hondeghem LM. Low safety index of domperidone: mechanism for increased odds ratio for sudden cardiac death. Acta Cardiol 2011 Aug;66(4):421-5.

[11] Santen G, van Zwet E, Danhof M, Della Pasqua O. From trial and error to trial simulation. Part 1: the importance of model-based drug development for antidepressant drugs. Clin Pharmacol Ther 2009 Sep;86(3):248-54.

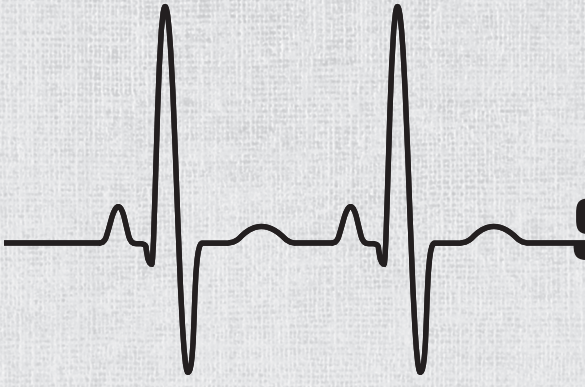
[12] Thompson SG, Pocock SJ. Can meta-analyses be trusted? Lancet 1991 Nov 2;338(8775):1127-30.

[13] Willan AR, Pinto EM. The value of information and optimal clinical trial design. Stat Med 2005 Jun 30;24(12):1791-806.



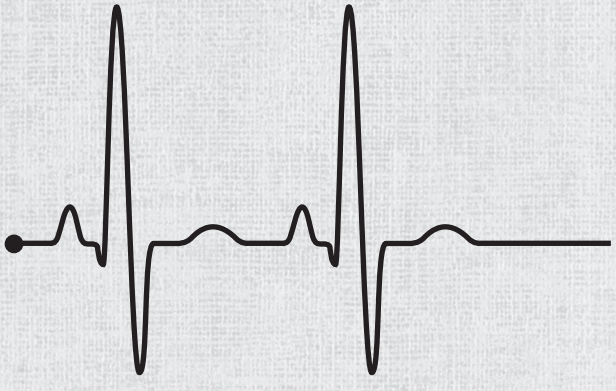
"Is it probable that probability brings certainty?"

Blaise Pascal , 1623 - 1662



Section II

Bayesian Methodology

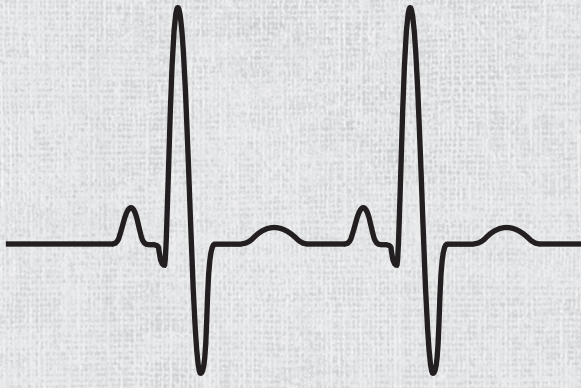


Assessing The Probability Of QTc-interval Prolongation In Clinical Drug Development

ASY Chain (1), KM Krudys (2), M Danhof (1), O Della Pasqua (1, 2)

(1) Leiden/Amsterdam Centre for Drug Research,
Division of Pharmacology, Leiden University, Leiden, The Netherlands
(2) Clinical Pharmacology and Discovery Medicine, GlaxoSmithKline,
Stockley Park, Uxbridge, United Kingdom

Clinical Pharmacology and Therapeutics, 2011 Dec;90(6):867-75



Chapter 3

ABSTRACT

Early in the course of clinical development of new non-antiarrhythmic drugs, it is important to assess the propensity of these drugs to prolong the QT/QTc-interval. The current regulatory guidelines suggest using the largest time-matched mean difference between drug and placebo (baseline-adjusted) groups over the sampling interval, thereby neglecting any potential exposure-effect relationship and nonlinearity in the underlying physiological fluctuation in QT values. Thus far, most of the attempted models for characterising drug-induced QTc-interval prolongation have disregarded the possibility of model parameterisation in terms of drug-specific and system-specific properties. Using a database consisting of three compounds with known dromotropic activity, we built a Bayesian hierarchical pharmacodynamic (PD) model to describe QT-interval, encompassing an individual correction factor for heart rate, an oscillatory component describing the circadian variation, and a truncated maximum-effect model to account for drug effect. The explicit description of the exposure-effect relationship, incorporating various sources of variability, offers advantages over the standard regulatory approach.

3.1 INTRODUCTION

The importance of evaluating possible effects of non-antiarrhythmic drugs on QT-interval has increased significantly in the past decade [1]. In fact, at present proarrhythmias due to drug-induced QTc prolongation are the second most common cause for withdrawal of a drug from the market [1, 2]. Such a cautionary measure derives from the notion that an increase in QTc-interval is associated with the risk of developing a potentially fatal arrhythmia, i.e., torsades de pointes [3-5]. Therefore, the ability to assess QT/QTc prolongation effects as early as possible in the course of clinical development of a new drug is helpful not only in understanding the potential risk-benefit ratio of a molecule but also in deciding whether to progress with the development of the compound, thereby avoiding unnecessary costs. To address some of the main issues related to cardiovascular safety, the most current regulatory guideline - the International Conference on Harmonisation (ICH) E14 document, published in 2005 - introduced the requirement to perform thorough-QT (TQT) studies [1, 6] as the basis for demonstrating a compound's liability to cause QTc prolongation. In addition to outlining the assessment procedures for evaluating prolonged ventricular repolarisation, the ICH E14 guidance defines a 10 msec increase in QTc-interval as a critical threshold for cardiovascular safety and requires the use of a positive control and supra-therapeutic doses of the investigational drug to ensure accuracy and sensitivity of the experimental protocol. Suggestions are also given regarding the timing of the studies as well as methodologies and interpretations used in the evaluation of QT measurements.

The primary analysis of a TQT study is based on the "double-delta" method, which is an assessment of the upper bound of the 95% one-sided confidence interval for the largest time-matched mean effect of the drug on the QTc-interval. The result of the assessment must exclude a 10 msec increase in QTc-interval if the drug is to be deemed safe, i.e., a negative study [1]. Requiring that the largest time-matched mean difference between the drug and placebo QTc-interval be ~5 msec or less implies that the one-sided 95% confidence interval should exclude an effect of $\geq$10msec for every single measurement. Although this is thought to be the most suitable approach, it can lead to multiplicity issues and positive bias [7, 8]. For instance, where electrocardiogram (ECG) measurements are taken at 10 post-dose sampling times, and in 9 of these, the one-sided 95% confidence intervals exclude a 10 msec while the tenth measurement has an upper limit of 11 msec, can one

conclude with certainty that the drug causes QT/QTc prolongation? To minimise the multiplicity issues arising from having to establish non-inferiority at each time point, investigators may choose to take measurements at only a bare minimum required number of time points. In other words, the more thorough approach of having ECG measurements taken at additional time points may itself carry a penalty; this is counterintuitive to the very purpose of TQTs.

From a methodological perspective, it should be noted that assessment of drug-induced QT-interval prolongation by non-antiarrhythmic compounds involves quantification of a weak signal and is prone to generation of false-positive and false-negative results. To overcome these issues and sustain enough power and acceptable Type I error rate, traditional statistical methods call for large study populations, representing considerable additional development costs. Another limitation of the proposed double-delta method is the inability to detect concentration-QTc relationships. The ICH guidance suggests that “establishing the relationship of drug concentrations to changes in QT/QTc-interval may provide additional information"; however, "this concept is not embedded into the statistical method. The suggestion is limited to the planning and interpretation of studies assessing cardiac repolarisation” [5]. In fact, Garnett et al. and many other authors have examined the potential role of the assessment of the drug concentration-QTc relationship and provided examples that illustrate how this type of assessment has been useful during regulatory review [9].

From a scientific perspective, the evidence gathered to date means that any "regulatory review of QT study is not complete without an assessment of concentration-QTc relationship” [10]. Assessing the relationship between the concentration of the drug and its effect on QT/QTc-interval can be achieved using alternative, yet equally powerful, quantitative methods; the latter would have a clear advantage over the standard E14 approach, namely the double-delta or any time point-based analysis. In particular, nonlinear mixed-effects modelling of the concentration-QT relationship allows the integration of data across all time points as well as across all available treatment groups. Moreover, it relies on individual responses, instead of averaging the QT response across time points, thereby enabling a better understanding of the uncertainty in response as well as the impact of outliers [10]. Despite the many modelling efforts attempted in the area of cardiovascular safety assessment [11-16], none has yielded a summary of the

risks associated with increasing exposures in such a way as to enable an easy comparison of compounds and to facilitate translational research and decision making. In addition, models often focus too much on drug-specific or experimental settings, overlooking the rationale underlying the assessment of causality in QTc prolongation. The model parameterisations proposed do not easily allow drug-specific properties to be distinguished from the biological or physiological system properties; nor do they facilitate translation from preclinical to clinical conditions.

Interestingly, the use of Bayesian hierarchical models has been rather limited in the field, as opposed to the more commonly-used frequentist approach in which maximum-likelihood estimation methods are used [17]. The Bayesian approach readily allows for the incorporation of prior knowledge that can exist for the system-related parameters such as QT/RR relationship and circadian variability. A wealth of knowledge has been generated throughout the history of cardiovascular safety assessment, and therefore the ability to incorporate prior information is invaluable. Another advantage to using Bayesian methods is that there is no requirement for linearity or normality in the data and therefore inference is based directly on the desired model rather than an approximation of it. Furthermore, the posterior distribution fully reflects all acknowledged sources of uncertainty and the fact that it is obtained in the form of a random sample means that numerous integrals of interest can be evaluated in a more straightforward manner. As a consequence, Bayesian methods enable one to express uncertainty about a parameter in terms of probability. It should also be noted that probability can be interpreted in a much more natural and transparent manner than is possible in frequentist terms. This transparency helps stakeholders in making decisions regarding the safety profile of a compound without the need to decipher frequentist-generated results.

Keeping in mind the aforementioned limitations with the current methods for the evaluation of QTc prolongation, the objective of this investigation was to demonstrate the feasibility of a Bayesian approach to characterise the exposure-effect relationship and the corresponding probability values of QTc-interval prolongation with respect to three compounds that are known to prolong the QT/QTc-interval (*d,l*-sotalol, moxifloxacin, and grepafloxacin), using clinical data from Phase I trials in healthy volunteers. Although any threshold values can be set, we have used a threshold of $\geq$ 10 msec in

assessing the probability curves for increases in QTc-interval as an illustration of the concept.

3.2 METHODS

Clinical data source

For this analysis *d,l*-sotalol, moxifloxacin and grepafloxacin - three compounds with known dromotropic activity - were selected from GlaxoSmithKline's clinical data repository. *d,l*-Sotalol treatment data were obtained from a three-way crossover, placebo-controlled (given twice), double-blinded, randomised trial design (study identifier EXP20001). Each subject was given a single oral dose of 160 mg *d,l*-sotalol or an equivalent dose of placebo. Moxifloxacin data were extracted from a two-way crossover, placebo-controlled, single–blinded trial aimed at evaluating the effect of repeated oral doses of lamotrigine on cardiac conduction (study SCA104648). Only data from sessions with placebo and moxifloxacin (400 mg), which was used as a positive control, was used for the purposes of our analysis [18]. Grepafloxacin data was obtained from a double-blind, placebo-controlled, multiple-dose crossover study design in which healthy female volunteers received a 600 mg dose of grepafloxacin or placebo (study GFX10901).

A summary of the relevant demographics, along with the time points for obtaining blood samples for pharmacokinetic and pharmacodynamic measurements pre- and post-dosing, pertaining to all three studies is shown in Table 1.

PK Sampling and Bioanalysis

For each of the compounds studied, whole blood (3 mL) was collected into plain tubes and left to clot at room temperature. After clot retraction occurred, the samples were centrifuged at 4°C at 1500 g – 2000 g for 10 minutes with the minimum of delay. Serum (~1 mL) was carefully pipetted into pre-labelled Nunc (Roskilde, Denmark) 1.8 mL polypropylene tubes and frozen (with the tubes in upright position) at -20°C until assay.

	d, l-Sotalol	Moxifloxacin	Grepafloxacin
Number of subjects	30	137	29
Gender	M: 18 (60%), F: 12 (40%)	M: 88 (64%), F: 49 (36%)	F: 29 (100%)
Age	mean = 36 (19-53)	mean = 27 (18-50)	mean = 24,2 (18-47)
Dose	placebo, 160 mg	placebo, 400 mg	placebo, 600 mg
PK sampling times	-30, -15, 5, 15, 30, 45 min, 1, 2, 4, 8, 10, 18, 24 h	-0.8, 0.35, 0.6, 0.85, 1, 1.3, 1.6, 2.1, 3.1, 4.1, 6.1, 12.1, 24.1 h	-0, 0,5, 1, 1.5, 2, 2,5, 3, 4, 6, 8, 12, 16, 24, 30, 36, 48 h
PD sampling times	-120, -105, -90, -75, -60, -45, -30, -15, 30 min, 1, 1.25, 2, 4, 6, 8, 10, 12, 18, 24 h	-1, -0.5, -0.1, 0.25, 0.5, 1, 1.5, 2, 2.5, 3, 4, 6, 8, 10, 12, 24 h	-25, -20, -15, -10, -5 min, 0.5, 1, 1.5, 2, 2,5, 3, 4, 6, 8, 12, 16, 24, 30, 36 h

Table 1. Clinical protocol and demographic details of the studies used in the analysis.

Concentrations of *d,l*-sotalol in serum were analysed using 96-well solid phase extraction/liquid chromatography-tandem mass spectrometry; the quantifiable range was 1.0 to 750 ng/mL with a coefficient of variation (CV) ranging from 5.86 – 13.29% and accuracy ranging from 93.95 – 109.9%.

Both moxifloxacin and grepafloxacin were analysed by means of high-performance liquid chromatography/tandem mass spectrometry, using a TurboIonspray interface (ABSCIEX, Foster City, CA) and multiple-reaction monitoring. Drugs were extracted from human plasma by solid-phase extraction using Oasis HLB cartridges (Waters, Milford, MA) with moxifloxacin-d4 as an internal standard. This method had a lower limit of quantification (LLQ) of 25.0 ng/mL and an upper limit of 5000 ng/mL. Quality-controlled results were deemed to be acceptable if they deviated by no more than 15% from the actual concentration. For the analytical run to be approved, no more than 33% of the results from quality control samples could be beyond ±15% of the actual concentration and at least 50% of the results for each concentration needed to be valid (not more than ±15% of the actual concentration).

ECG measurements

All subjects were required to stay overnight at the study location before receiving the relevant study drug. The drug was administered with 200 mL of water. The consumption of all foods and beverages was restricted in accordance with the respective protocols so as to minimise random variability in ECG measurements. The subjects were also asked to refrain from strenuous physical activity before and during the trial period so that ECG measurements would not be altered by external factors.

Twelve-lead ECG machines (Marquette™, GE Healthcare, Bucks, UK) were used for the assessment of QT, QTc[Bazett], QTc[Fridericia], RR, and HR. Strict procedures were followed to ensure that subjects remained in the supine position at least 5 min before ECG recordings commenced and throughout the duration of the recordings. Triplicate ECG printouts were obtained at each sampling time. QT-intervals were measured manually from lead II by a cardiographist who was blinded with regard to the treatment randomisation. Subsequently, the average was taken from the available replicates at each sampling time. Pre-dose ECG measurements were used as baseline values for each study. PK and PD sampling time points for each study are shown separately in Table 1 since the sampling schedules in each trial were very different.

Pharmacokinetics (PK) modelling

Population pharmacokinetic models were built for each compound using nonlinear mixed-effects modelling techniques as implemented in NONMEM VI (v 1.0, ICON Development Solutions, Ellicott City, MD, USA) [19]. Model selection was based on the likelihood ratio test, parameter point estimates and their respective confidence intervals as well as parameter correlations and goodness-of-fit plots. For the likelihood ratio test, the significance level was set at $p=0.05$, after the inclusion of one parameter; this corresponds to a decrease of 3.84 points in the minimum value of the objective function (MVOF) under the assumption that the difference in MVOF between two nested models is chi-square distributed. Fixed and random effects were introduced into the model in a step-wise fashion. Inter-individual variability in PK parameters was assumed to be log-normally distributed. Different residual variability models describing measurement and model errors were tested to achieve the best fit. The final model was used to generate

individual plasma concentrations at each of the sampling times associated with ECG measurements.

Pharmacodynamic-pharmacodynamic (PKPD) modelling

The pharmacodynamic (PD) model describing QTc-interval was developed in a modular way that allowed a distinction to be made between physiological (QT and RR) and drug-specific (concentration) parameters; an individual correction factor for RR-interval (heart rate), an oscillatory component describing the circadian variation and a truncated Emax model to capture drug effect [20, 21] were used, as presented in Equation1:

$$QT_c = QT_0 \cdot RR^{\alpha} + A \cdot \cos\left(\frac{2\pi}{24}(t - \phi)\right) + slope \cdot C$$

Equation 1

where, QT_0 [msec] is the intercept of the QT-RR relationship. The sex of the patient was included as a covariate for this parameter whenever applicable. RR [sec] is the interval between successive R waves, α is the individual heart rate correction factor, A [msec] is the amplitude of circadian rhythm, *t* is the clock time, φ is the phase, slope [msec/concentration] is the linear pharmacodynamic relationship, and C is the predicted concentration of drug at the time points of QT-interval measurements.

Instead of applying traditional maximum likelihood methods to model QTc-interval, a Bayesian adaptation of the model is proposed using WinBUGS 1.4 [22, 23]. The model was implemented in a hierarchical manner, in which subject-specific levels (between-subject variability) were considered for all parameters (i.e., α , A , φ and slope). On the basis of physiological considerations, these parameters can be assumed to be independent and uncorrelated. Furthermore, given that each drug relates to a different drug, study was not included as a term in the analysis.

The use of a Bayesian approach ensures systematic translation of the concentration-effect relationship and interpretation of the risk associated with drug-specific effects. Each parameter is estimated as a mean value, with its precision and random effects describing the between-subject variability. An example of the model code is available in the Appendix. Among other

advantages, Monte-Carlo Markov-Chain (MCMC) methods allow incorporation of prior information during parameter estimation. In addition, they yield full posterior distributions from which 95% credible intervals can be derived, thereby facilitating understanding of the model and parameter uncertainty. Detailed information on Bayesian analyses of PKPD models may be found elsewhere [24, 25].

Priors

The k^{th} observed QT measurement for the j^{th} occasion for the i^{th} (QT_{ijk}) individual was assumed to be normally distributed around the individual predicted QT measurement f_{ijk} with an unknown precision τ:

$$QT_{ijk} \sim N(f_{ijk}, \tau), \quad \tau = \frac{1}{\sigma^2}$$

Non-informative priors were specified as:

$$\theta \sim MVN(\bar{\mu}, \Sigma^{-1}), \quad \bar{\mu} = \begin{bmatrix} QT_0 \\ \alpha \\ A \\ \phi \\ Slope \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \Sigma^{-1} = 0.0001 \cdot I_5$$

where θ is a vector of population mean parameter estimates, $\bar{\mu}$ is the prior of the population means, Σ^{-1} is the precision of the prior for the population mean parameter values and *I* is the identity matrix. The inverse of the between-subject variance, Ω^{-1}, arises from a Wishart distribution: $\Omega^{-1} \sim W(\rho\Omega, \rho)$, with $\rho=5$ degrees of freedom, where Ω represents our prior guess at the order of magnitude of the covariance matrix. Finally, non-informative Gamma (0.0001, 0.0001) priors were assumed for measurement precision and inter-occasion variability of QT_0.

Goodness- of- fit and modelling diagnostics

Two MCMC chains were run in parallel for 25,000 samples and pooled together to provide final parameter estimates. The posterior distributions provided by WinBUGS enabled the model-predicted outcomes to be compared with the observed values. Appropriateness of parametric distributions as well as the drug model was assessed using the deviance information criterion (DIC) as a measure of the goodness-of-fit [26, 27]. The number of iterations was deliberately chosen to exclude any residual correlation. Convergence was determined using the Gelman-Rubin test statistic [28-30] and also visually, using history and autocorrelation plots.

Probability of QTc-interval prolongation ≥ 10 msec

In conjunction with model parameter estimation, a step function was set to calculate the probability of observing QTc-interval prolongation ≥ 10 msec at peak concentrations (Cmax). The procedure consisted in obtaining the cumulative distribution of values in the posterior samples of the product of Slope and Cmax which were ≥10 msec.

3.3 RESULTS

Pharmacokinetic modelling

Each compound was modelled separately; the details are shown in Table 2. Despite the use of different parameterisations for volume of distribution, we found that a two compartment pharmacokinetic model best described the time course of concentrations in plasma for all three QTc-prolonging drugs. For *d,l*-sotalol and grepafloxacin, the absorption phase was characterised by a lag time, whereas for moxifloxacin, a combination of zero- and first-order absorption processes was required. In the case of *d,l*-sotalol, it was found that body weight had a significant effect on the volume of distribution and that creatinine clearance affected clearance. None of the covariates was found to affect the pharmacokinetics of moxifloxacin and grepafloxacin. In addition to standard statistical criteria, goodness-of-fit plots were created so as to visually inspect the fit for each of the models (Figure 1). The maximum population-predicted concentration (Cmax) were 932 ng/mL for *d,l*-sotalol,

1771 ng/mL for moxifloxacin and 2515 ng/mL for grepafloxacin. A summary of the estimated model parameters is shown in Table 2.

Given that sample collection for the drug concentration quantifications was carried out at time points different from those for ECG measurements, the purpose of the pharmacokinetic models was to allow imputation of drug concentrations at the time point of each ECG measurement so that pair-wise concentration-effect data would be available for pharmacodynamic (PD) modelling.

Compound	***d,l*-Sotalol**		**Moxifloxacin**		**Grepafloxacin**	
Number of compartments	2		2		2	
Parameter estimates	Mean	IIV	Mean	IIV	Mean	IIV
Ka [/h]	1.27	1.39	2.21	0.786	1.32	0.779
D1 [/h]	-	-	0.629	0.725	-	-
CL [L/h]	9.39	1.03	13.4	0.0161	0.0823	0.0065
Vss [L]	147	1.02	-	-	0.217	0.776
V2 [L]	-	-	122	0.0828	-	-
V3 [L]	-	-	55.4	0.196	-	-
ALAG [h]	0.45	1.16	-	-	0.413	-
Q [/h]	5.58	0	78.4	-	-	-
Error model used	proportional		proportional		Proportional	
Covariates	Creatine on CL Weight on Vss		N/A		N/A	

Table 2. Summary of pharmacokinetic parameter estimates for d,l-sotalol, moxifloxacin, and grepafloxacin.

Pharmacokinetic-pharmacodynamic modelling

The pharmacokinetic-pharmacodynamic (PKPD) model allowed the system-specific parameters (i.e., those related to the individual correction for the QT-RR relationship and circadian rhythm) to be estimated separately from the drug-specific parameters (i.e., drug-induced QTc-interval prolongation). Accordingly, it was found that QT_0, the exponent of the individual QT-RR

correction factor (α), the amplitude (A) and the phase (φ) of the 24-hour circadian rhythm were in the same range of values for all three compounds. The one parameter that is noticeably different between the compounds was the slope value describing the relationship between drug concentration and QTc-interval. As expected, this effect is drug-specific. Estimates for all parameters used in the PKPD model are presented in Table 3, including the between-subject variability (BSV) and inter-occasion variability (IOV).

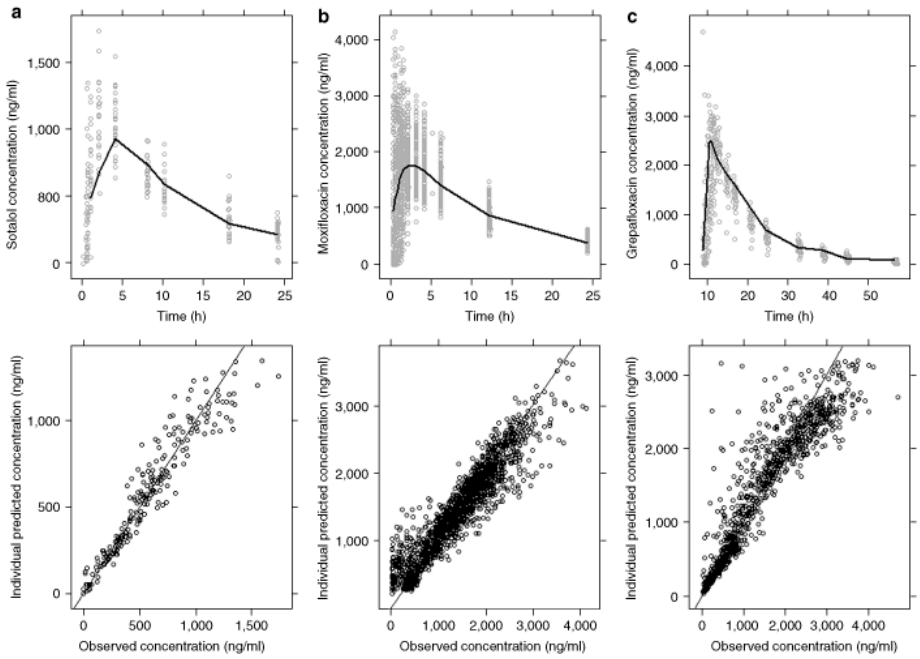


Figure 1. Predicted (solid line) and observed (dots) concentration time profiles for (a) d,l-sotalol, (b) moxifloxacin and (c) grepafloxacin. The lower panel shows the goodness-of-fit plots for the individual predicted vs. observed drug concentrations.

In addition to the statistical criteria, accuracy of model predictions was assessed graphically by superimposing population-predicted QT-interval for each compound onto observed QT values. A direct comparison between model predictions and the observed QTc-intervals was not possible because Bazett's correction factor was applied to the original data. Goodness-of-fit plots showing individual predictions vs. observed values confirm the absence of bias (Figure 2). Likewise, diagnostic plots show that the model accurately described the PKPD relationship between QTc-interval and predicted concentrations of *d,l*-sotalol, moxifloxacin and grepafloxacin (Figure 3).

Using Bayesian inference, the probability of QTc-interval prolongation ≥10 msec was calculated based on the cumulative distribution of values in the posterior samples of the difference between the QTc–interval values at baseline and after drug administration. This procedure enables the characterisation of drug effect on QTc-interval throughout the observed concentration range and beyond, providing a systematic translation of model findings into clinically relevant measures. Figure 4 shows the mean drug effect along with the 95% credible intervals throughout a wide range of concentration values for *d,l*-sotalol, moxifloxacin and grepafloxacin. The 95% credible intervals, which are analogous to confidence intervals in frequentist statistics, provide a measure of uncertainty in the PKPD relationship. With therapeutic doses (160 mg for *d,l*-sotalol, 400 mg for moxifloxacin, and 600 mg for grepafloxacin), the probability of having a QT prolongation ≥10 msec at observed maximum predicted concentration range was 100% for all three compounds.

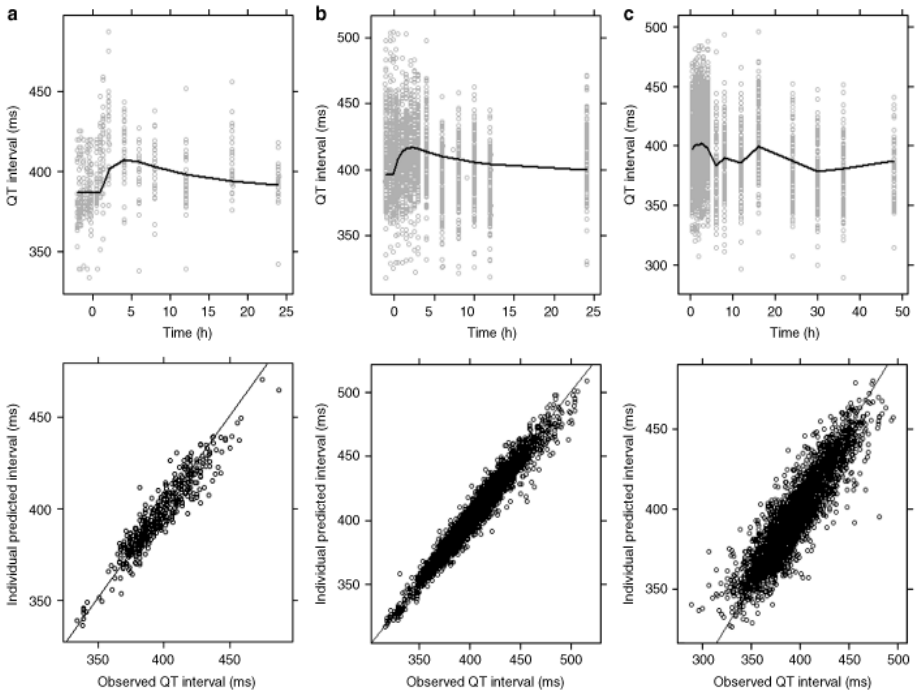


Figure 2. Predicted (solid line) and observed (dots) QT-interval vs. time following administration of (a) d,l-sotalol, (b) moxifloxacin and (c) grepafloxacin. The lower panel shows the goodness-of-fit plots for the individual predicted vs. observed QT-interval.

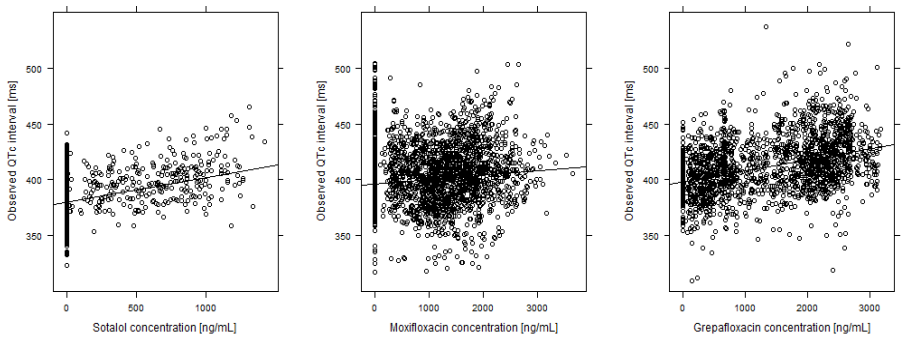


Figure 3. Pharmacokinetic-pharmacodynamic relationship between QTc-interval and plasma concentrations of (left) d,l-sotalol, (middle) moxifloxacin and (right) grepafloxacin. The solid line represents the regression for the population predicted slope and intercept. Dots show the observed QTc-interval and the corresponding (predicted) individual concentration data

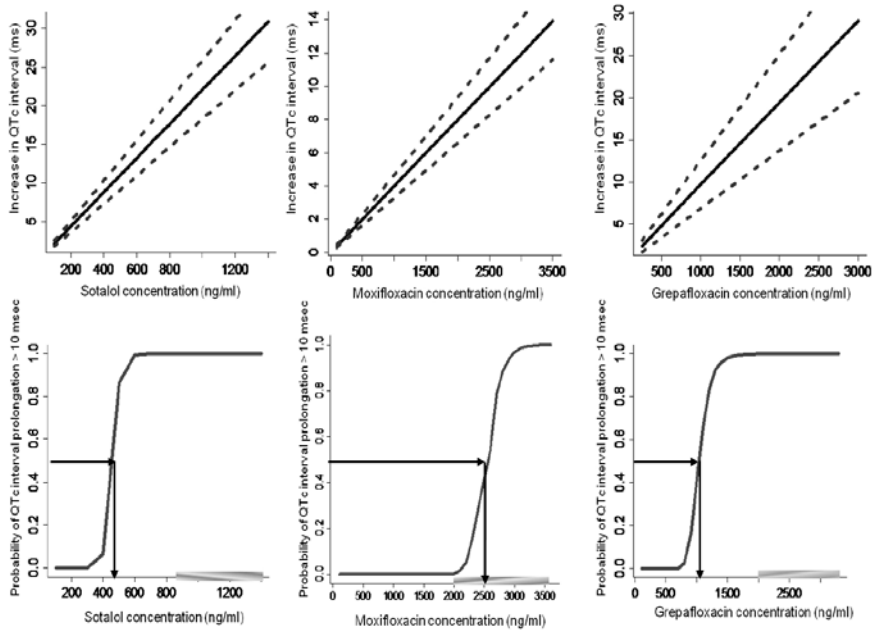


Figure 4. Pharmacokinetic-pharmacodynamic relationships (mean and 95% CI) and corresponding probability curves for QTc-interval prolongation > 10 msec vs. the predicted plasma concentrations of d,l-sotalol (left), moxifloxacin (middle) and grepafloxacin (right).

	***d,l*-Sotalol (n=30)**	**Moxifloxacin (n=137)**	**Grepafloxacin (n=29)**
QT0 [msec]	380 (375-383)	396 (394-398)	393 (388-398)
Sex Effect [msec]	17 (7-27)	8 (5-12)	-
A	0.28 (0.25-0.31)	0.41 (0.40-0.43)	0.29 (0.25-0.32)
A [msec]	3.4 (2.7-4.1)	2.7 (2.2-3.0)	5.1 (3.9-6.1)
φ [h]	6.4 (5.4-7.4)	9.5 (7.7-11.1)	5.8 (4.6-7.1)
Slope [msec/conc]	0.022 (0.018-0.026)	0.0040 (0.0033-0.0046)	0.0098 (0.0069-0.013)
IOV (QT0) %	1.3 (1.1-1.7)	0.65 (0.54-0.79)	2.5 (2.0-3.4)
BSV (QT0) %	2.2 (1.6-3.1)	1.9 (1.5-2.3)	2.3 (1.6-3.4)
BSV (α) %	18 (13-26)	15 (11-19)	21 (15-31)
BSV (A) %	3.6 (2.2-9.2)	4.4 (2.9-9.6)	2.4 (1.5-9.5)
BSV (φ) %	12 (7-28)	14 (7-39)	16 (8-42)
BSV (Slope) %	34 (26-47)	83 (79-96)	65 (52-87)
Residual Error [msec]	5.9 (5.7-6.18)	5.3 (5.2-5.4)	15 (14.7-15.4)
Probability Effect ≥ 10 msec at Cmax	1.00	1.00	1.00

*Mean values with (95% confidence interval)

Table 3. Mean parameter estimates of the Bayesian pharmacokinetic-pharmacodynamic model for d,l-sotalol and moxifloxacin, and grepafloxacin with 95% confidence intervals.

3.4 DISCUSSION

The E14 document recommends that almost all drugs, especially non-antiarrhythmic compounds, undergo careful clinical testing in TQT study. The underlying concept is that a study in healthy volunteers using supra-therapeutic doses of a drug is sufficient to detect small increases in QTc-interval. This finding can unveil potentially larger prodromic effects in patients, who are more prone to develop pro-arrhythmias [31, 32]. However, one of the main limitations of this approach is that it does not explicitly account for individual plasma drug concentration levels and the corresponding PKPD relationships when assessing treatment effect, yielding biased estimates of the true drug effect [33]. Antiarrhythmic and non-antiarrhythmic drugs that prolong the QTc-interval have been shown to do so in an exposure- or concentration-dependent manner [34-37]. Moreover, other physiological and protocol design factors contribute to changes in QTc-interval, and these might not be mitigated by randomisation methods and other statistical techniques.

To address this limitation, we aimed in this study to show how a Bayesian hierarchical model can be used to describe the relationship between drug concentration and QTc-interval prolongation with respect to *d,l*-sotalol, moxifloxacin and grepafloxacin. The use of an integrated modelling approach has also allowed explicit identification of drug-induced component of QTc prolongation, discriminating it from other components such as variation in heart rate and circadian rhythm. It should be noted that the drug-specific parameter in this model (slope) does not independently indicate whether there are risks associated with the use of the compound, nor does it indicate whether the development of the drug should be stopped. Rather, it provides an unambiguous description of the PKPD relationship, taking into account various sources of variability. Furthermore, the use of a linear or log-linear relationship to describe drug-specific effects offers advantages as compared with the standard regulatory approach in that it yields estimates of the exposure required to reach clinically relevant increase in QT-interval without the requirement for parameterisation of maximum effect (E_{max}) [38]. The resulting PKPD relationships are also similar to values that have been previously reported using other methodologies. For example, Dixon *et al.* found an increase in QTc of ~6 msec per 1000 ng/mL QT_c [18], and Bloomfield reported 3.9 msec/1000 ng/mL [39]; our finding was 4 msec per 1000 ng/mL.

From a drug development perspective, this approach is highly relevant because 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. A link model can be introduced to infer effect-site concentrations without the burden of subsequent drug-specific model adaptations. The proposed Bayesian analysis provides clearly interpretable clinical measurements which may aid in the decision-making 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. These concepts contrast with the current recommendations by the ICH E14, which suggest the use of supra-therapeutic doses in TQT studies, without the incorporation of pharmacokinetic information. Unfortunately, many investigators do not realise that, as a consequence of such recommendations, the effect size observed under these conditions cannot properly be deduced for therapeutic doses. Similar arguments may also be applicable to situations to which exposure could potentially be altered because of other factors that are relevant for the assessment of the benefit:risk ratio, such as drug-drug interactions or comorbidity conditions. The ability to extrapolate the effect size across different dose ranges and to estimate the associated risk levels can greatly enhance the decision making process for regulators, clinicians and drug developers. In this investigation, we have shown the probability values associated with increases in QTc-interval $\geq$10 msec, over a given range of concentrations. However, this threshold can easily be changed to higher or lower values to accommodate other types of assessments, rendering this model useful for different types of analysis or scenarios, including the comparison of drug effects across species.

From a statistical perspective, whether our proposed methodology overcomes the high rate of type I error observed with the traditional approach remains to be further investigated. Nevertheless, we expect that the use of model-based approaches will be a requirement for the assessment of drug safety. Accurate judgement of the risk and liability associated with a novel chemical or biological entity cannot be performed without distinguishing between drug-specific parameters and the so-called system-specific parameters. This concept has been elegantly illustrated by Danhof et al. and proposed as the basis for translational pharmacology research, particularly at the interface between drug discovery and early drug

development [40]. In this sense, we envisage the use of our model as a framework for clinical trial simulations and optimisation of clinical protocol design.

The high attrition rates in drug discovery and development are largely attributable to issues of efficacy and/or safety. Regulatory guidelines should ensure that the appropriate level of evidence is available for better decision-making *before* the approval process. This requirement does not seem to be warranted by the current ICH E14. Concerted efforts are needed to ensure that patients are granted access to drugs on the basis of scientific evidence of safety. The bias imposed by traditional methods should not continue to be ignored.

REFERENCES

[1] Guidance for Industry: E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrythmic Potential for Non-Antiarrythmic Drugs. In: Services USDoHaH, editor; 2005.

[2] Darpo B, Nebout T, Sager PT. Clinical evaluation of QT/QTc prolongation and proarrhythmic potential for nonantiarrhythmic drugs: the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use E14 guideline. Journal of Clinical Pharmacology 2006;46(5):498-507.

[3] Hutmacher MM, Chapel S, Agin MA, Fleishaker JC, Lalonde RL. Performance characteristics for some typical QT study designs under the ICH E-14 guidance. Journal of Clinical Pharmacology 2008;48(2):215-24.

[4] Pollard CE, Valentin JP, Hammond TG. Strategies to reduce the risk of drug-induced QT interval prolongation: a pharmaceutical company perspective. Br J Pharmacol 2008 Aug;154(7):1538-43.

[5] Zareba W. Drug induced QT prolongation. Cardiol J 2007;14(6):523-33.

[6] Interdisciplinary Review Team for QT Studies. Manual of policies and procedures MAPP 6020.14. Division of Cardiovascular and Renal Products, Center for Drug Evaluation and Research, U.S. Food and Drug Administration(2007).

[7] Tsong Y, Shen, M., Zhong, J., Zhang, J. Statistical Issues of QT Prolongation Assessment Based on Linear Concentration Modeling. Journal of Biophamaceutical Statistics 2010(18):564-.

[8] Boos DD, Hoffman D, Kringle R, Zhang J. New confidence bounds for QT studies. Stat Med 2007 Sep 10;26(20):3801-17.

[9] Garnett CE, Beasley N, Bhattaram VA, Jadhav PR, Madabushi R, Stockbridge N, et al. Concentration-QT relationships play a key role in the evaluation of proarrhythmic risk during regulatory review. Journal of Clinical Pharmacology 2008;48(1):13-8.

[10] Bloomfield D, Krishna R. Commentary on the clinical relevance of concentration/QTc relationships for new drug candidates. Journal of Clinical Pharmacology 2008;48(1):6-8.

[11] Malik M, Andreas JO, Hnatkova K, Hoeckendorff J, Cawello W, Middle M, et al. Thorough QT/QTc study in patients with advanced Parkinson's disease: cardiac safety of rotigotine. Clin Pharmacol Ther 2008 Nov;84(5):595-603.

[12] Malik M, Hnatkova K, Ford J, Madge D. Near-thorough QT study as part of a first-in-man study. J Clin Pharmacol 2008 Oct;48(10):1146-57.

[13] Rohatagi S, Carrothers TJ, Kuwabara-Wagg J, Khariton T. Is a thorough QTc study necessary? The role of modeling and simulation in evaluating the QTc prolongation potential of drugs. J Clin Pharmacol 2009 Nov;49(11):1284-96.

[14] Vandemeulebroecke M, Lembcke J, Wiesinger H, Sittner W, Lindemann S. Assessment of QT(c)-prolonging potential of BX471 in healthy volunteers. A 'thorough QT study' following ICH E14 using various QT(c) correction methods. Br J Clin Pharmacol 2009 Sep;68(3):435-46.

[15] Pater C. Methodological considerations in the design of trials for safety assessment of new drugs and chemical entities. Curr Control Trials Cardiovasc Med 2005 Feb 3;6(1):1.

[16] Anand SP, Ghosh, S.K. A Novel Bayesian Approach to Assessing the Risk of QT Prolongation. Institute of Statistics Mimeo Serios #2620. North Carolina.

[17] Kathman SJ, Williams DH, Hodge JP, Dar M. A Bayesian population PK-PD model of ispinesib-induced myelosuppression. Clin Pharmacol Ther 2007 Jan;81(1):88-94.

[18] Dixon R, Job S, Oliver R, Tompson D, Wright JG, Maltby K, et al. Lamotrigine does not prolong QTc in a thorough QT/QTc study in healthy subjects. Br J Clin Pharmacol 2008 Sep;66(3):396-404.

[19] Beal S, Sheiner, L. NONMEM user's guide. San Francisco: NONMEM project group; 1999.

[20] Bachman WJ, Gillespie, W.R. Truncated sigmoid Emax models: a reparameterization of the sigmoid Emax model for use wtih truncated PK/PD data. American Society for Clinical Pharmacology adn Therapeutics Meeting; 1998; 1998.

[21] Piotrovsky V. Pharmacokinetic-pharmacodynamic modeling in the data analysis and interpretation of drug-induced QT/QTc prolongation. AAPS J 2005;7(3):E609-24.

[22] Spiegelhalter D, Thomas, A., Best, N., Gilks, W. BUGS: Bayesian Inference using Gibbs Sampling, Version 1.4.1 Cambridge: MRC Biostatistics Unit; 2003.

[23] Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS - A Bayesian modelling framework: concepts, structure, and extensibility. Stat Comput 2000(10):325-37.

[24] Lunn DJ, Best N, Thomas A, Wakefield J, Spiegelhalter D. Bayesian analysis of population PK/PD models: general concepts and software. J Pharmacokinet Pharmacodyn 2002 Jun;29(3):271-307.

[25] Duffull SB, Kirkpatrick CM, Green B, Holford NH. Analysis of population pharmacokinetic data using NONMEM and WinBUGS. J Biopharm Stat 2005;15(1):53-73.

[26] Spiegelhalter D AK, Myles J. Bayesian approaches to clinical trials and health care evaluations. New York: Wiley; 2004.

[27] Spiegelhalter D, Best, N., Carlin, B. Bayesian measures of model complexity and fit. J Roy Stat Soc Ser B 2002(64):583-639.

[28] Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. J Comput Graph Stat 1998(7):434-55.

[29] Gelman A, Rubin, D. Inference from iterative simulation using multiple sequences. Stat Sci 1992(7):457-511.

[30] Gelman A, Carlin, J.B., Stern, H.S., Rubin, D.B. Bayesian Data Analysis,. London: Chapman & Hall; 2003.

[31] Wheeler, W. Evolution of the thorough QT/QTc study. Regulatory Rapporteur 2010 7(11), 21-23.

[32] Shah RR. Drug-induced prolongation of the QT interval: regulatory dilemmas and implications for approval and labelling of a new chemical entity. Fundam Clin Pharmacol 2002 Apr;16(2):147-56.

[33] Chapel S. HMM, Bockbrader H., de Greef R., Lalonde R.L. 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. Clinical Pharmacology & Therapeutics 2010;89(1):75-80.

[34] Allen MJ, Nichols DJ, Oliver SD. The pharmacokinetics and pharmacodynamics of oral dofetilide after twice daily and three times daily dosing. Br J Clin Pharmacol 2000 Sep;50(3):247-53.

[35] Lin YL, Chan, K. Pharmacokinetic and pharmacodynamic characterization of non-antiarrhythmic QT-prolonging drugs associated wtih torsade de pointes. Drug Information Journal 2008;42:211-9.

[36] Barbey JT, Sale ME, Woosley RL, Shi J, Melikian AP, Hinderling PH. Pharmacokinetic, pharmacodynamic, and safety evaluation of an accelerated dose titration regimen of d,l-sotalol in healthy middle-aged subjects. Clin Pharmacol Ther 1999 Jul;66(1):91-9.

[37] Jonker DM, Kenna LA, Leishman D, Wallis R, Milligan PA, Jonsson EN. A pharmacokinetic-pharmacodynamic model for the quantitative prediction of dofetilide clinical QT prolongation from human ether-a-go-go-related gene current inhibition data. Clin Pharmacol Ther 2005 Jun;77(6):572-82.

[38] Schoemaker RC, van Gerven JM, Cohen AF. Estimating potency for the Emax-model without attaining maximal effects. J Pharmacokinet Biopharm 1998 Oct;26(5):581-93.

[39] Bloomfield DM, Kost JT, Ghosh K, Hreniuk D, Hickey LA, Guitierrez MJ. The effect of moxifloxacin on QTc and implications for the design of thorough QT studies. Clin Pharmacol Ther 2008 Oct;84(4):475-80.

[40] Danhof M, de Lange EC, Della Pasqua OE, Ploeger BA, Voskuyl RA. Mechanism-based pharmacokinetic-pharmacodynamic (PK-PD) modeling in translational drug research. Trends Pharmacol Sci 2008 Apr;29(4):186-91.

APPENDIX - EXAMPLE OF THE WINBUGS CODE REQUIRED FOR MODELLING QTC-INTERVAL PROLONGATION

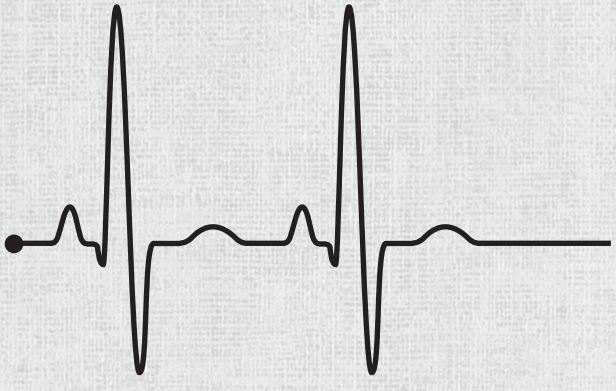
```
model{
for(i in 1:N) {
    for(k in 1:P){
        for(j in offset[P*(i-1)+k]:(offset[P*(i-1)+k + 1] -
1)){
QT[j] ~ dnorm(mu[j],tau.pop)
mu[j]<-(1+exp(theta[i,2])*cos((2*3.1416/24)*(CTIME[j]-
        exp(theta[i,3])))+(0.00001*theta[i,4])*
        C[j])*(beta[i,k]) *pow(RR[j],exp(theta[i,1]))
QTC[j]<-QT[j]/pow(RR[j],exp(theta[i,1]))-
exp(theta[i,2])*cos((2*3.1416/24)*(CTIME[j]-
exp(theta[i,3])))
res[j]<-(QT[j]-mu[j])/sigma2
P.res[j]<-phi(res[j])                #p-value for res
Y.pred[j] ~ dnorm(mu[j],tau.pop)     #predicted response
P.pred[j]<-step(QT[j]-Y.pred[j])     #indicator of whether
                                     #observed >predicted
}#data

beta[i,k] ~ dnorm(mu.beta[i],beta.tau)

}#period

theta[i,1:5] ~ dmnorm(mu.theta[i,1:5],R[ , ])
mu.theta[i,1]<- fixed.effect[1]
mu.theta[i,2]<-fixed.effect[2]
mu.theta[i,3]<-fixed.effect[3]
mu.theta[i,4]<-fixed.effect[4]
mu.theta[i,5]<-fixed.effect[5]*
  (1 - SEX[i])+fixed.effect[6]*SEX[i]
mu.beta[i]<-exp((theta[i,5]))
}#subject

fixed.effect[1:6] ~ dmnorm(mean[1:6], prec[1:6 ,1:6 ])
R[1:5,1:5] ~ dwish(omega[,],5)
tau.pop ~ dgamma(1.0E-3,1.0E-3)
sigma2<-1.0/sqrt(tau.pop)
beta.tau ~ dgamma(0.001,0.001)
}
```



Establishing *In Vitro* To Clinical Correlations In Evaluation Of Cardiovascular Safety Pharmacology

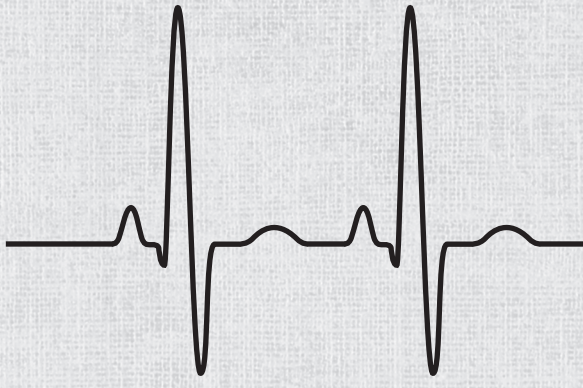
ASY Chain (1)*, VFS Dubois (1)*, M Danhof (1) MCJM Sturkenboom (2, 3), OE Della Pasqua (1, 4) on behalf of the Cardiovascular Safety Project Team, TI Pharma PKPD Platform.

(1) Leiden/Amsterdam Centre for Drug Research,
Division of Pharmacology, Leiden University, Leiden, The Netherlands
(2) Department of Medical Informatics and Epidemiology,
Erasmus Medical Centre, The Netherlands
(3) Department of Epidemiology, Erasmus Medical Centre,
The Netherlands
(4) Clinical Pharmacology and Discovery Medicine, GlaxoSmithKline,
Stockley Park, Uxbridge, United Kingdom

Members of the TI Pharma Cardiovascular Safety Project
involved in the preparation of this manuscript:

Eli Lilly & Co: Dinesh De Alwis, Derek Leishman
GlaxoSmithKline: Jackie Bloomer, Nick McMahon, Phil Milliken
Johnson & Johnson: David Gallagher, An Vermeulen
Pfizer: Piet van der Graaf, Mark Holbrook
Erasmus Medical Centre: Jeanne Dieleman, Miriam Sturkenboom

Submitted for publication



Chapter 4

ABSTRACT

Aims: Given the similarities in QTc response between canines and humans, dogs are often used in preclinical cardiovascular safety studies. The objective of our investigation was to characterise the PKPD relationships and translational gaps across species following the administration of three compounds known to cause QTc-interval prolongation, namely cisapride, d,l-sotolol and moxifloxacin.

Methods: Pharmacokinetic and pharmacodynamic data from experiments in conscious dogs and clinical trials following administration of the drugs were used. First, pharmacokinetic modelling and deconvolution methods were applied to derive drug concentrations at the time of each QT measurement. A Bayesian PKPD model was then used to describe QT prolongation, allowing discrimination of drug-specific effects from other physiological factors known to alter QT-interval duration. A threshold of $\geq$ 10 msec was used to explore the probability of prolongation following drug administration.

Results: A linear correlation was found between the PKPD relationships in dogs and humans. The drug-specific parameter (slope) in dogs and humans were respectively, 0.0045 vs. 0.009, 0.002 vs. 0.021, 0.00056 vs. 0.0039 msec/nM for cisapride, d,l-sotalol and moxifloxacin. Despite the statistically significant differences between species, QTc prolongation $\geq$ 10 msec at therapeutic exposure range in humans could be predicted from data in conscious dogs.

Conclusions: Our findings indicate that the slope of PKPD relationship in conscious dogs may be used as basis for the prediction of drug-induced QTc prolongation in humans. Furthermore, the risk for QTc prolongation can be expressed in terms of the probability associated with an increase $\geq$ 10 msec, allowing direct inferences about the clinical relevance of the pro-arrhythmic potential of a molecule.

4.1 INTRODUCTION

Cardiovascular safety issues are one of the main causes of failure and attrition in drug discovery and development [1, 2]. One challenge has been the possibility to translate pre-clinical findings and anticipate the implications for further clinical development. A major pre-requisite in translational research remains however, the characterisation of concentration-effect relationships, which can be used as basis for extrapolation of pharmacokinetic and pharmacodynamic data across species [3]. The aforementioned pre-requisite has been overlooked by the focus on compliance to regulatory guidelines, which currently do not endorse the need for the generation of safety data in an integrated manner, in which drug exposure and/or relevant biomarkers of drug effect are analysed using a model-based approach [2, 4, 5].

One area which can benefit substantially from such an approach is the evaluation of the propensity of non-antiarrhythmic drugs in prolonging QT/QTc-interval [2, 6]. Prolongation of the QT-interval has been associated with the development of polymorphic ventricular tachycardia or torsade de pointes (TdP), one of the major reasons for post-market withdrawal [1, 7]. Despite the fact that QT-interval can easily be measured as part of routine electrocardiogram (ECG) monitoring, assessment of drug-induced QTc-interval prolongation relies on empirical criteria, preventing the accurate scaling of drug effects across species [8].

Currently, various *in vitro* and *in vivo* systems are available for assessing QT prolongation [9-11], which support decision-making for the progression of assets into clinical development. These assays provide in most cases qualitative measures of drug effect and must therefore be complemented by safety data in a thorough-QT (TQT) trial [2, 12-14]. In contrast, the use of a model-based approach offers a strictly quantitative basis for translational research, which can enhance the integration and extrapolation of preclinical data, eventually dismissing the need for additional evidence from a TQT trial.

Recently, we have proposed a Bayesian pharmacokinetic-pharmacodynamic model to assess the probability of QTc-interval prolongation of three compounds known to prolong the QT/QTc-interval (*d,l*-sotalol, moxifloxacin, grepafloxacin) using data from Phase I trials in healthy volunteers [15]. Our method relies on model parameterisation that allows drug-specific

properties to be distinguished from the biological or physiological system properties, facilitating the translation of data from preclinical to clinical conditions. In addition, the 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. It also offers the possibility to estimate posterior parameter distributions, which fully reflect all acknowledged sources of uncertainty. From a drug development perspective, this approach can be 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 objective of this investigation was therefore to demonstrate its value for predicting the clinical effects of cisapride, *d,l*-sotalol and moxifloxacin based on the exposure-effect relationships and the corresponding probability of QTc-interval prolongation from *in vivo* dog data. Even though any threshold values can be defined for probability curves, our analysis assumes comparable effect size for the increase in QTc-interval in dogs and in humans (i.e. $\geq$ 10 msec).

4.2 METHODS

4.2.1. EXPERIMENTAL DATA

Preclinical protocols

A four-way crossover design was used for the evaluation of QTc-interval in conscious, freely moving beagle dogs. All dogs were chronically instrumented with radio telemetry probes measuring blood pressure (BP), ECG and body temperature (T). The ECG electrodes were placed in a lead II position. Animals were administered vehicle, a sub-therapeutic, a therapeutic and a supra-therapeutic oral dose of each compound.

Blood sampling scheme for pharmacokinetics was based on preliminary pharmacokinetic data to ensure accurate characterisation of absorption, distribution and elimination phases. ECG was monitored continuously over

the period of 24 hours, with samples collected every 30 sec. Further details on the experimental design and data acquisition are summarised in Table 1a.

	Cisapride	***d,l*-Sotalol**	**Moxifloxacin**
Number of animals	8	6	8
Dose [mg/kg]	Vehicle, 0.6, 2, 6	Vehicle, 4, 8	Prestudy, vehicle, 3, 10, 30
PK sampling times [h]	0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 5, 6, 8, 24	0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 6, 8, 24	0, 0.5, 1, 2, 4, 8, 24, 36, 48
PD sampling times	Every 30 sec, averages over 24 h	Every 5 min, averages over 24 h	Every 1 min, averages over 24 h
PK parameter, vital signs, demographic covariates, ECG parameters	Plasma drug concentration, clock time, Heart rate, Blood pressure, Weight, Sex, LvPr, QT, RR, QRS		

Table 1a. Treatment, measurement variables and sampling scheme details for the pre-clinical experiments in conscious dogs.

Clinical protocols

All three protocols were Phase I studies in healthy volunteers. For cisapride, a randomised, placebo-controlled, dose-escalating design was used in which subjects received up to five doses. Due to safety issues, the study was not continued after the fourth dose. *d,l*-Sotalol data was extracted from a double-blind, randomised, placebo-controlled, three-way crossover design in which each subject received one active treatment and two placebo doses. Data for moxifloxacin was available from the positive control arm of a two-way crossover, single-blind, randomised, placebo-controlled trial aimed at evaluating the effect of repeated oral doses of lamotrigine on cardiac conduction. Further details on the experimental design and data acquisition are summarised in Table 1b.

	Cisapride	*d, l-sotalol*	Moxifloxacin
Number of subjects	24	30	137
Gender	M: 14 (58%) F: 10 (42%)	M: 18 (60%) F: 12 (40%)	M: 88 (64%) F: 49 (36%)
Age [yr]	mean = 40 (19-54)	mean = 36 (19-53)	mean = 27 (18-50)
Dose [mg]	placebo, 10, 20, 40, 80	placebo, 160	placebo, 400
PK sampling times [h]	0, 1, 1.5, 2, 3, 4, 6, 12, 24	-0.5, -25, 0.83, 0.25, 0.5, 0.75, 1, 2, 4, 8, 10, 18, 24	-1, -0.5, -0.83, 0.25, 0.5, 1, 1.5, 2, 2.5, 3, 4, 6, 8, 10, 12, 24
PD sampling times [h]	-24, -23, -22.5, -21, -20, -18, -12, 0, 1, 1.5, 2, 3, 4, 6, 12, 24	-2, -1.75, -1.5, -1, -0.75, -0.5, -0.25, 0.5, 1, 1.25, 2, 4, 6, 8, 10, 12, 18, 24	-1, -0.5, -0.83, 0.25, 0.5, 1, 1.5, 2, 2.5, 3, 4, 6, 8, 10, 12, 24
PK parameter, vital signs, demographic covariates, ECG parameters	Plasma drug concentration, clock time, Heart rate, Weight, Sex, QT, QTc_b, QTc_f, RR		

Table 1b. Treatment, measurement variables and sampling scheme details for the clinical studies in healthy volunteers.

4.2.2. PHARMACOKINETIC SAMPLING PROCEDURES AND BIO-ANALYSIS

Preclinical experiments

Cisapride - Approximately 0.5 mL samples were obtained from the jugular vein and transferred into uniquely labeled EDTA tubes. All samples were mixed gently and placed on crushed water-ice until centrifugation at 3000 g for 5 minutes, which was carried out at approximately 4 ˚C. The resultant plasma was separated, transferred to uniquely labeled clear matrix tubes, frozen immediately over solid carbon dioxide and stored at approximately –80 ˚C or below until analysis. Cisapride concentrations were analysed using a method based upon protein precipitation followed by LC-MS/MS analysis. The lower limit of quantification (LLQ) was 0.5 ng/mL for a 25 μL aliquot of dog plasma.

Sotalol - Approximately 0.6mL blood samples were obtained from the jugular vein and transferred to BD-microtainers K2E with K2EDTA, and centrifuged at 8000 rcf for 3 min in a cooled centrifuge. The resultant plasma was separated, transferred to uniquely labeled clear matrix tubes and stored at approximately –18 ˚C or below until analysis. Racemic *d,l*-sotalol was analysed using a LC–MS/MS method (Waters UPLC) and an API 4000 mass spectrometer. An Acquity UPLC® BEH hilic, 1.7 µm: 2.1x100mm column was used for the separation by the LC method (with A = 0.1% Formic Acid and B = CH3OH). The lower limit of quantification (LLQ) was 1 ng/mL.

Moxifloxacin - Approximately 1.0 mL whole blood samples were obtained from the jugular vein and transferred into uniquely labelled potassium EDTA tubes where plasma was separated by centrifugation. Serum concentrations of moxifloxacin (Ryan Scientific, Mt. Pleasant, SC) were determined by acetonitrile precipitation, followed by analysis using the LC–MS/MS method. Samples were prepared by an addition of 600 µL of acetonitrile containing 500 ng/mL of lomefloxacin (Sigma Chemical Co. Ltd., St. Louis, MO) as an internal standard to 25 µL of serum. Moxifloxacin serum concentrations were determined by an Applied Biosystems API4000 triple quadrupole mass spectrometer (Applied Biosystems, Foster City, CA, USA) equipped with a TurboV™ source. Spectra was acquired by multiple reaction monitoring (MRM) in the positive ionisation mode. A calibration curve was constructed using peak area ratios of moxifloxacin and the internal standard calibration samples of lomefloxacin, with a weighted (1/x2) linear least squares regression analysis.

The lower limit of quantification (LLQ) was 1ng/mL and the upper limit was 5000 ng/mL. The same procedure was used for two separate analysis days.

Clinical protocols

For all bio-analysis pertaining to the clinical studies, 3 mL blood samples were collected into vacutainers for the determination of serum concentrations of the specific compound. Blood samples were left to clot at room temperature. After clot retraction occurred, samples were centrifuged at 4°C at 1500 g for 10 minutes with the minimum of delay. Serum (approximately 1 mL) was carefully pipetted into pre-labelled Nunc 1.8 mL polypropylene tubes, then frozen in an upright position at -20°C until assay.

All samples were kept frozen in Cardice pellets during transit and transport to various analysis facilities.

Cisapride - Plasma concentrations of cisapride were determined by a validated high performance liquid chromatography (HPLC) method. The analytical range was between 5.00 and 500 ng/mL. Each analytical batch of cisapride assays consisted of a set of 8 calibrators, freshly prepared in blank heparinised plasma. The lowest calibrator was at the level of the lower limit of quantification and the highest calibrator was at the level of the upper limit of quantification. Back-calculated values of the calibration standards had to be within 15% validation-criterion (20% at LLOQ). The coefficent of variance (CV) per calibration level never exceeded the validation-criterion of $\leq$ 15% CV (at LLOQ $\leq$ 20%) and at 1.6% maximum.

Sotalol - Serum *d,l*-sotalol concentrations were analysed using 96-well solid phase extraction/liquid chromatography-tandem mass spectrometry with a quantifiable range of 1.0 to 750 ng/mL with % CV ranging from 5.86 – 13.29 and % accuracy ranging from 93.95 – 109.9 %.

Moxifloxacin - Moxifloxacin was extracted from human plasma by solid-phase extraction using Oasis HLB cartridges with moxifloxacin-d4 as an internal standard. Extracts are analysed by HPLC-MS/MS using a TurboIonspray interface and multiple reaction monitoring. This method has a lower limit of quantification (LLQ) of 25.0 ng/mL using a 50 µL aliquot of human plasma. The analytical range was 25.0 – 5000 ng/mL. A calibration standard was omitted if the back-calculated concentration deviates from actual by more than 20%. Within and between-run precision, and accuracy of quantification during the validation was better than 9.62% across the concentration range.

4.2.3. ECG MONITORING AND SAMPLING

Preclinical experiments

The ECG waveforms were captured with DSI implants, with EMKA IOX (EMKA technologies, France). The data were recorded throughout the experiments. An EMKA ECG Auto analysis software (IOX EMKA technologies, France) was used for the analysis of cisapride, while the Notocord (Notocord-hem, KRN42a, Notocord, France) was used for *d,l*-sotalol and the

PONEMAH (Cardiovascular Toxicology Laboratory, Covance Laboratories, Greenfield, USA) was used for moxifloxacin. ECG scans were averaged every 30 sec (cisapride), 1min (moxifloxacin), or 5 min (*d,l*-sotalol) and individually corrected for RR. The QT-intervals were capped if there was more than 10% difference between consecutive QT measurements. This was done to prevent telemetry equipment noise affecting the assessment of drug-induced effects on QT-interval. Capping of the data occurred mainly during feeding time (6h after dosing).

Given the limited number of sampling points required for modeling purposes, somewhat sparse sampling times were subsequently extracted from the original recordings. Sampling windows were selected according the pharmacokinetic properties of each compound, in such a way that data were collected more frequently during the absorption phase and during peak concentrations. Each data point used in the pharmacokinetic-pharmacodynamic analysis reflected the average of QT and RR observed during the corresponding sampling interval. In brief, the following sampling intervals were used:

Cisapride - 0-2 h after dose: 2 min averaged intervals, 2-10 h after dose: 5 min averaged intervals, 10-24 h after dose: 15 min averaged intervals.

Sotalol - 0-10 h after dose: every 5 min averaged, and 10-24 h after dose: every 15 min averaged

Moxifloxacin - 0-2 h after dose: 2 min averaged intervals, 2-10 h after dose: 5 min average samples, 10- 47 h after dose: 15 min averaged intervals.

Clinical protocols

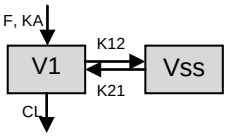
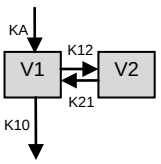
ECG monitoring was performed with 12-lead electrocardiogram, using Marquette ECG machines measuring QT, QTc[Bazett], QTc[Fridericia], RR, and HR for three compounds. The time used to reference the ECG measurements were actual clock-time as recorded in the data collection forms, rather than the planned sampling time indicated in the study protocol. For each study, subjects were kept in a supine position while ECG recordings were made.

In general, subjects in all three studies were restricted to the type and amount of beverages allowed during the study period. Food in predefined

amount was served at strict times of the day. Subjects were also asked to refrain from strenuous exercises.

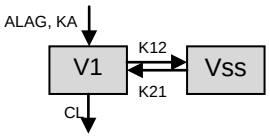
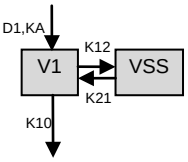
4.2.4. PHARMACOKINETIC (PK) MODELLING

Time matched concentration and QT-interval values were required for the assessment of the pharmacokinetic-pharmacodynamic relationships. Therefore, pharmacokinetic modelling and deconvolution techniques were used to generate the corresponding concentration values for each ECG sampling time. Pharmacokinetic modelling was performed using non-linear mixed effects techniques in NONMEM VI (ICON, Maryland, USA) for the clinical data and the *d,l*-sotalol preclinical data. For the preclinical pharmacokinetic modelling of moxifloxacin, NONMEM 7.1.2 (ICON, Maryland, USA) was used. Deconvolution was the method of choice for the description of the cisapride data in dogs and was performed using deconvolution and noncompartmental analysis methods in WINNONLIN v4.1. (Pharsight, USA). Models used and parameters estimated are shown in Tables 2a and b for the three different compounds for dogs and human respectively.

	Cisapride	***d, l*-Sotalol**	**Moxifloxacin**
Graphical representation of the PK model	*NA*		
Parameters estimated	AUC**, T1/2**	F*, Ka, V1, Vss, Cl, Q	Ka, V1, V2, Q, Cl, F*
BSV	NA	V1, Vss, Cl, Q	Ka, V1, V2, Q, Cl
Covariates	NA	NA	NA
Derived PK timepoints	1-2hr: every 2 mins, 2-10hr: every 5 mins, 10-24hr every 15 mins	1-10hr: every 5 mins, 10-24hr: every 15 mins	1-2hr: every 2 mins, 2-10hr: every 5 mins 10- 47hr: every 15 mins

**F estimated separately per dose level **Derived using non-compartmental analysis*

Table 2a. PK models and parameters estimated for preclinical protocols

	Cisapride*	*d, l*-Sotalol	Moxifloxacin
Graphical representation of the PK model	NA		
Parameters estimated	CL, Vc, Vp, Q, Ka, D0, ALAG, F	Ka, ALAG, CL, V, Q	D1, Ka, CL, V2, V3, Q
Between subject variabilities	CL, Vc, Vp, Q, Ka, D0, ALAG, F	Ka, ALAG, CL, V, Q	D1, Ka, CL, V2, V3, Q
Covariates	N/A	Body weight, creatinine clearances	N/A

**Denotes parameters as reported in final clinical study report.*

Table 2b. PK models used and parameters estimated for clinical protocols.

4.2.5. PHARMACOKINETIC-PHARMACODYNAMIC (PKPD) MODELLING

Model building was performed in WinBUGS version 1.4.2 [16-19]. While the same model describing QTc-interval was created for all three compounds for both species, each analysis was conducted independently. A Bayesian adaptation of a model comprising of three components: an individual correction factor for RR-interval (heart rate), an oscillatory component describing the circadian variation and a truncated Emax model to capture drug effects [13, 20], was used as presented in Equation1:

$$QT = QT_0 \cdot RR^{\alpha} + A \cdot \cos\left(\frac{2\pi}{24}(t-\phi)\right) + slope \cdot C$$

Equation 1

where, QT_0 [msec] is the intercept of the QT-RR relationship (sex was included as a covariate for this parameter), RR [sec] is the interval between successive R waves, α is the individual heart rate correction factor, A [msec] is the amplitude of circadian rhythm, t is the clock time, φ is the phase, slope

[msec/concentration] is the linear pharmacodynamic relationship, and C is the predicted concentration of drug at the time of QT measurements.

Priors

The k^{th} observed QT measurement for the j^{th} occasion for the i^{th} (QT_{ijk}) individual was assumed to be normally distributed around the individual predicted Qt measurement f_{ijk} with an unknown precision τ:

$$QT_{ijk} \sim N(f_{ijk}, \tau), \quad \tau = \frac{1}{\sigma^2}$$

Non-informative priors were specified as:

$$\theta \sim MVN(\bar{\mu}, \Sigma^{-1}), \quad \bar{\mu} = \begin{bmatrix} QT_0 \\ \alpha \\ A \\ \phi \\ Slope \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \Sigma^{-1} = 0.0001 \cdot I_5$$

where θ is a vector of population mean parameter estimates, $\bar{\mu}$ is the prior of the population means, Σ^{-1} is the precision of the prior for the population mean parameter values and *I* is the identity matrix. The inverse of the between subject variance, Ω^{-1}, arises from a Wishart distribution: $\Omega^{-1} \sim W(\rho\Omega, \rho)$, with ρ=5 degrees of freedom, where Ω represents our prior guess at the order of magnitude of the covariance matrix. Finally, non-informative Gamma (0.0001, 0.0001) priors were assumed for measurement precision and inter-occasion variability of QT_0.

Goodness -of-fit and modelling diagnostics

Two MCMC chains were run for 25 000 samples and pooled to provide parameter estimates. The posterior distributions provided by WinBUGS allowed the model-predicted outcomes to be compared with the observed

values. Appropriateness of parametric distributions as well as the drug model was assessed using the deviance information criterion (DIC) as a measure of the goodness-of-fit [21, 22]. The number of iterations was deliberately chosen to exclude any residual correlation. Further details of the model have been described elsewhere [15].

The R package R2WinBUGS was used to execute WinBUGS while running a session in R (The R project - http://www.r-project.org/). Convergence was assessed visually by monitoring the dynamic traces of Gibbs iterations and by computing the Gelman–Rubin test statistic [22-25].

4.2.6. PROBABILITY OF QTC-INTERVAL PROLONGATION

The analysis was performed with a step function in WinBUGS 1.4.3 using the slope and an inter-individual correction factor for gender differences (see Equation 2) at the following concentrations: 50, 200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, 2500, 3000, 3500, 4000, 6000, 9000, 10000, 15000, 30000nM. Data was expressed in nano-molar (nM) to allow direct comparison of the different compounds. And the concentration steps were chosen to cover the sigmoid curve in an appropriate fashion,

$$\text{Probability of } \geq 10\,\text{msec prolongation (at C)} = \text{step}(0.00001^{\text{F(Gender)}} \cdot \text{Slope} \cdot \frac{10}{\text{C}})$$

Equation 2

where, C is the drug concentration, and slope is the QT increase per unit drug concentration.

Values were plotted for all three compounds to compare the clinical and preclinical probabilities of drug induced QTc prolongation. By using the nano-molar unit, the potency of increasing QTc between compounds can be compared and used to create an interspecies correlation.

To compare if there was statistically significant differences between the QTc prolongation per concentration unit in the different compounds and species, we performed Z-score tests, where the the means and standard deviations of the parameter estimates of the apriory distributions were used for the calculations.

4.3 RESULTS

Pharmacokinetic (PK) models

Given that most PK observations were at different time points compared to the ECG measurements or in a disproportionate fashion with the abundance of ECG recordings in dogs, it was necessary to perform PK modeling and simulation to obtain a balanced dataset for the subsequent PKPD analysis using the Bayesian approach. Each compound was modelled separately for each species, with exception of cisapride for which modelling attempts yielded unsatisfactory results in dogs, while time-matched pairs of concentration and ECG measurements were available in humans. Given the variability in the cisapride data and the scope of our analysis, a deconvolution method was applied in dog data instead. A summary for the parameter estiamtes are shown in Tables 3a and b for dog and humans respectively. Model performance is depicted in Figures 1b and 1c for *d,l*-sotalol and moxifloxacin. Only the observed data and population averages are shown for cisapride in Figure 1a. Details of the pharmacokinetic models are found in Tables 2a and b for dogs and humans respectively.

Compound	**Cisapride**		***d, l*-Sotalol**		**Moxifloxacin**	
Parameter estimates	Median	Min, Max	Mean	IIV%	Mean	IIV%
AUC 0.6 mg/kg	794	295, 1548	-	-	-	-
AUC 2.0 mg/kg	6977	860, 12453				
AUC 6.0 mg/kg	6873	971, 17361				
T1/2	4.82	3.57, 7.57	-	-	-	-
Ka [/h]	-	-	1.41	-	1.78	-
CL [L/h]	-	-	2.79	18.3	3.39	21
Vss [L]	-	-	-	-	43.23	-
V2 [L]	-	-	3.15	100	-	-
V3 [L]	-	-	21.8	15.8	-	-
ALAG [h]	-	-	1.41	-	-	-
Q [/h]	-	-	16.5	35.4	-	-
F*	-	-	1, 121, 29.7	-		-

**F estimated separately per dose level*

Table 3a. Summary of pharmacokinetic parameter estimates for cisapride, d,l-sotalol, and moxifloxacin using preclinical protocols.

Compound	Cisapride*		*d, l*-Sotalol		Moxifloxacin	
Parameter estimates	Median	min, max	Mean	IIV%	Mean	IIV%
Ka [/h]	-	-	1.27	57.36	2.21	88.65
D0 [/h]	-	-	-	-	0.629	85.14
CL [L/h]	18	7.7, 33.1	9.39	17.15	13.4	12.69
Vss [L]	-	-	147	12.73	-	-
V2 [L]	173.5	17.8, 255.1	-	-	122	28.77
V3 [L]	265.8	12.42, 1384	-	-	55.4	44.27
ALAG [h]	-	-	0.45	38.47	-	-
Q [/h]	34.1	10.1, 395.1	5.58	0	78.4	-
F	0.75	0.47, 1	-	-	-	-

**Values for 20 mg dose.*

Table 3b. Summary of pharmacokinetic parameter estimates for cisapride, d,l-sotalol, and moxifloxacin using clinical protocols.

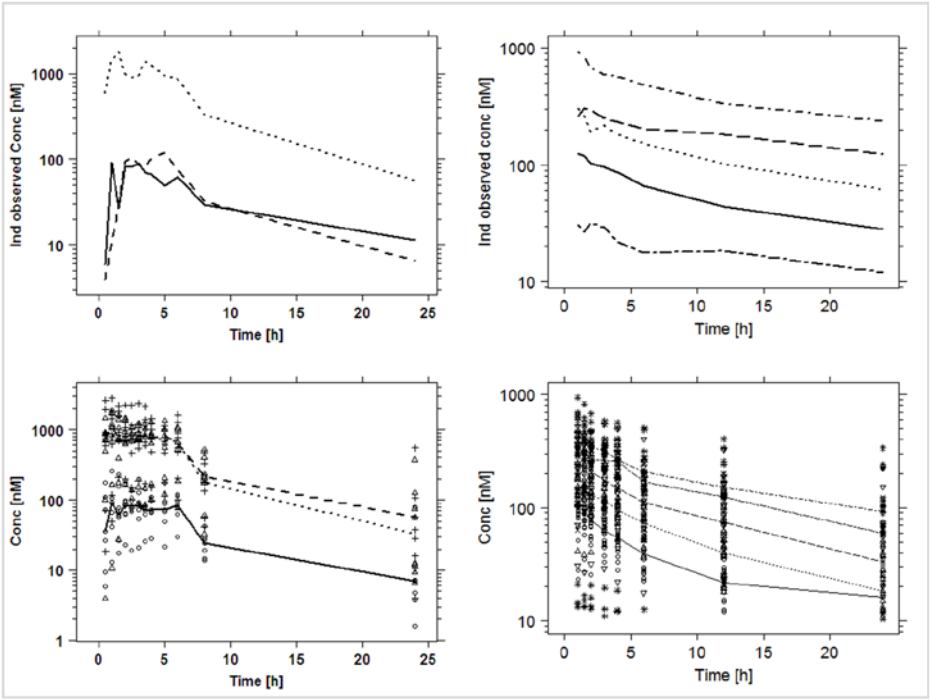


Figure 1a. Observed cisapride concentrations vs. time of a typical subject (upper panel) and observed population median concentrations vs. time (lower panel) for dogs (left) and humans (right).

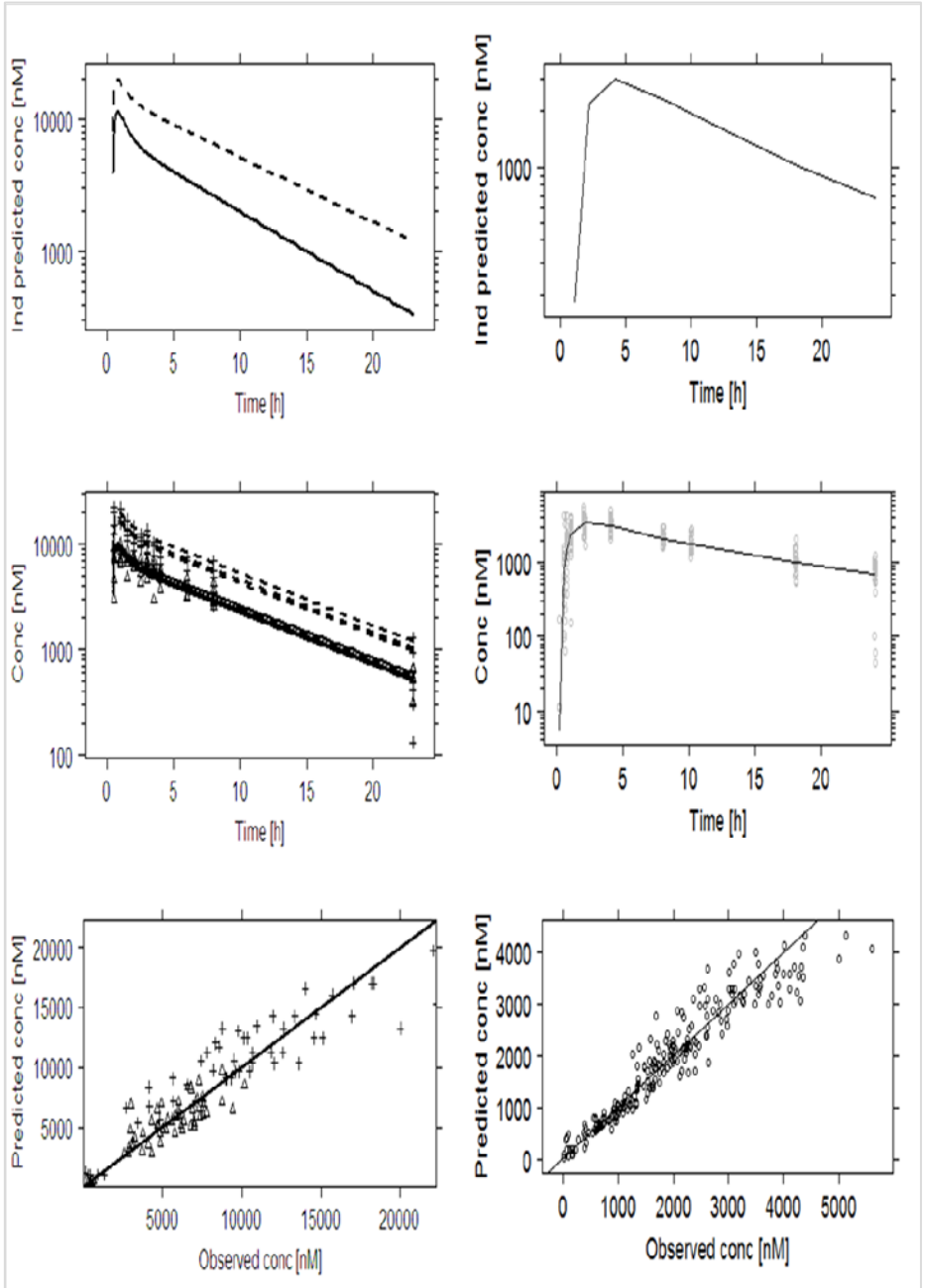


Figure 1b. Individual predicted d,l-sotalol concentration vs. time (top panels), population predicted d,l-sotalol concentration vs. time (middle panels) and individual predicted d,l-sotalol concentrations vs. observed d,l-sotalol concentrations (lower panels) for dogs (left) and humans (right).

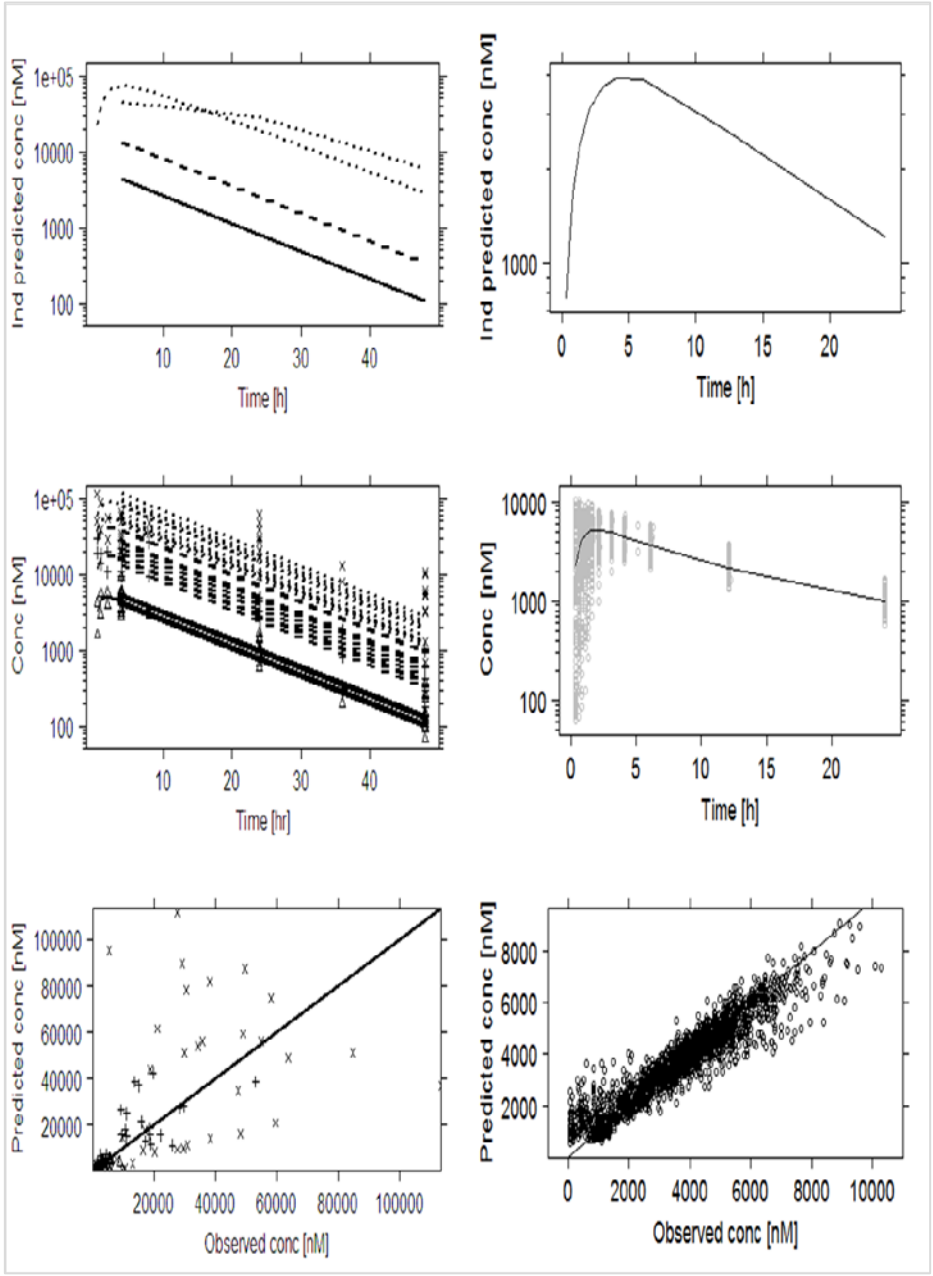


Figure 1c. Individual predicted moxifloxacin concentration vs. time (top panels), population predicted moxifloxacin concentration vs. time (middle panels) and individual predicted moxifloxacin concentrations vs. observed moxifloxacin concentrations (lower panels) for dogs (left) and humans (right).

Pharmacokinetic-pharmacodynamic (PD) modelling

PKPD modelling was preformed using the recorded ECG measurements and drug concentrations at the corresponding sampling time. If not available at the exact sampling times, concentrations were imputed based on simulations and deconvolution techniques, as indicated above. Given the overwhelming amount of ECG recordings in dogs, data filtering was applied to ensure data sets of comparable sizes were used in the joint analysis with clinical data. Data filtering was performed taking the absorption and disposition profiles into account, so that peak and elimination phase were accurately represented. The filtering scheme has been summarised in Table 1a.

System-specific parameters, i.e. baseline QT (QT0), the QT-RR correction factor (α), the amplitude (A) and phase (Φ), all showed values within the same range within each species. The main difference was in the drug-specific parameter (slope) between compounds and also between dogs and humans. A summary of the analysis for each compound is presented in Tables 3a-c. In addition to the typical Bayesian criteria for parameter convergence and model acceptance statistical criteria , graphical summaries using goodness-of-fit plots (i.e., observed vs. predicted QT values) per compound and species are shown in Figures 2a-c.

Parameters	Dogs Total Drug (n = 7)	Human (n = 24)
QT_0 [msec]	238 (237 - 240)	386 (382 - 390)
α	0.26 (0.20 - 0.33)	0.18 (0.14 - 0.23)
A [msec]	5.6 (3.9 - 8.1)	3.3 (1.1 - 5.95)
φ [h]	19.9 (15.8 - 25.8)	4.3 (2.4 - 8.68)
Slope [msec/conc]	0.0045 (0.00096 - 0.0098)	0.09 (0.087 - 0.12)
BSV (QT_0) %	6.49 (6.46 - 6.48)	2.23 (1.75 - 3.01)
BSV (α) %	89 (47-190)	17.31 (11.48 - 26.95)
BSV (A) %	7.14 (5.32 - 10.93)	6.52 (3.73 - 14.57)
BSV (φ) %	7.08 (4.40 - 13.13)	19.04 (11.79 - 35.3)
BSV (Slope) %	105 (66 - 180)	48.2 (21.94 - 93.5)
Residual Error [msec]	5.8 (5.7 - 5.9)	11.66 (11.13 - 13.04)
Prob Effect ≥ 10 msec at Cmax	0.75	1.0
Cmax [nM]	2808	936

Table 4a. Mean parameter estimates of the Bayesian pharmacokinetic-pharmacodynamic model for cisapride for dogs and human with 95% confidence intervals.

Parameters	Dogs Total Drug (n = 5)	Human (n = 30)
QT_0 [msec]	255 (253 - 257)	387 (383 - 392)
α	0.18 (0.11 - 0.3)	0.27 (0.24 - 0.3)
A [msec]	6.58 (2.09 - 19.78)	3.3 (2.4 - 4.29)
φ [h]	12.15 (6.51 - 35.77)	6.22 (5.05 - 7.58)
Slope [msec/conc]	0.002 (0.0006 - 0.008)	0.021 (0.017 - 0.026)
BSV (QT_0) %	6.25 (6.23 - 6.28)	5.08 (5.05 - 5.11)
BSV (α) %	137 (55-410)	80 (54 -123)
BSV (A) %	11.02 (8.9 - 17.3)	6.38 (4.08 - 10.15)
BSV (φ) %	11.11 (4.73 - 30.70)	8.44 (13.98 - 24.24)
BSV (Slope) %	36.8 (22.6 - 71.1)	108 (49.7 - 18.2)
Residual Error [msec]	9.0 (8.8 - 9.3)	5.9 (5.68 - 6.14)
Prob Effect ≥ 10 msec at Cmax	0.9	1.0
Cmax [nM]	22310	5605

Table 4b. Mean parameter estimates of the Bayesian pharmacokinetic-pharmacodynamic model for d,l-sotalol for dogs and human with 95% confidence intervals.

Parameters	Dogs Total Drug (n = 8)	Human (n = 137)
QT_0 [msec]	240 (238 - 242)	399 (394 - 403)
α	0.28 (0.22 - 0.35)	0.40 (0.38 - 0.42)
A [msec]	4.6 (3.1 - 7.0)	2.4 (1.7 - 2.9)
φ [h]	23.1 (15.1 - 34.6)	10.0 (7.3 - 12.9)
Slope [msec/conc]	0.00056 (0.00002 - 0.0014)	0.0039 (0.0033 - 0.0044)
BSV (QT_0) %	6.46 (6.43 - 6.48)	5.01 (4.98 - 5.04)
BSV (α) %	86 (48 - 177)	41 (33 - 52)
BSV (A) %	9.8 (7.4 - 14.8)	5.3 (3.6 - 8.0)
BSV (φ) %	13 (8 - 23)	18 (10 - 31)
BSV (Slope) %	25 (17 - 39)	41 (29 - 53)
Residual Error [msec]	9.4 (9.3 - 9.5)	5.3 (5.2 - 5.4)
Prob Effect ≥ 10 msec at Cmax	1.0	1.0
Cmax [nM]	112930	10300

Table 4c. Mean parameter estimates of the Bayesian pharmacokinetic-pharmacodynamic model for moxifloxacin for dogs and human with 95% confidence intervals.

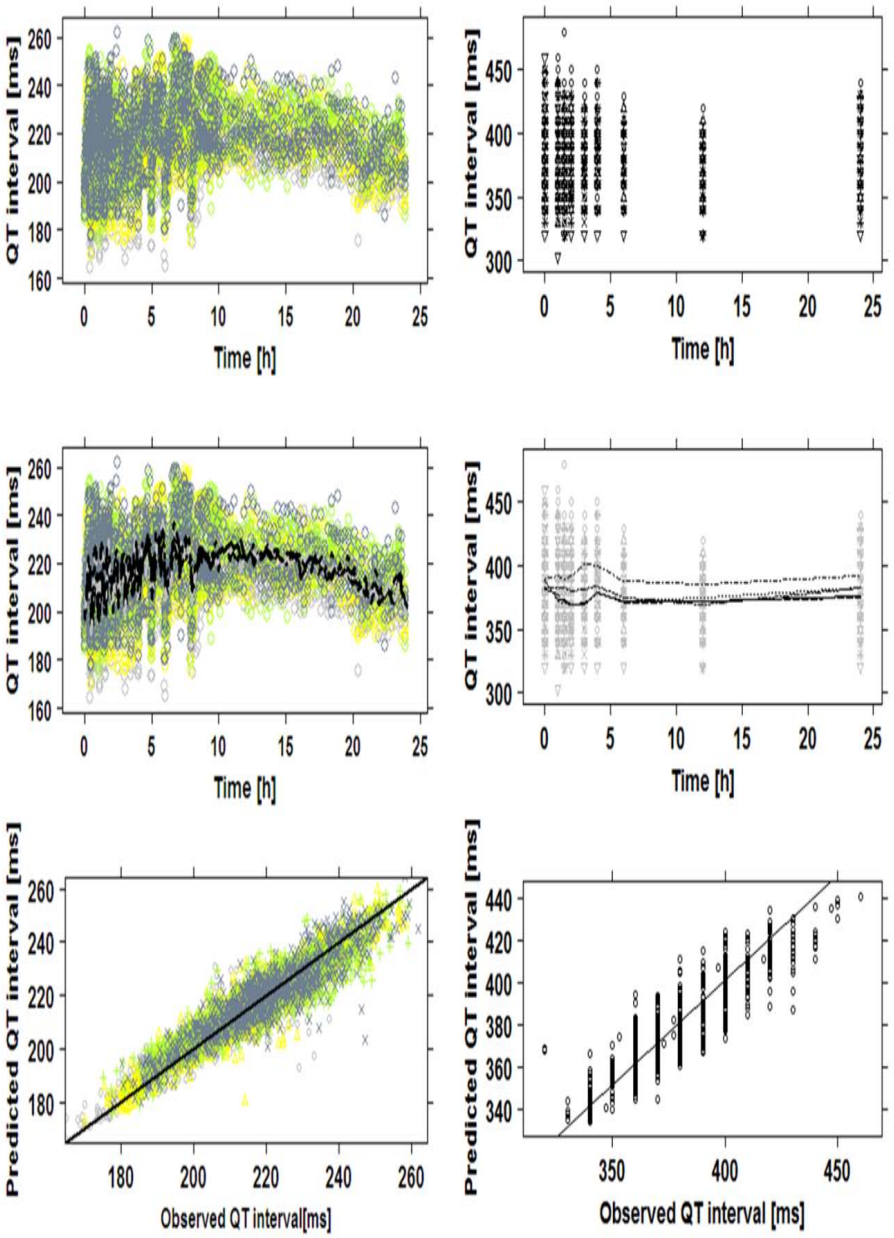


Figure 2a. Pharmacokinetic-pharmacodynamic relationship between QTc-interval and plasma concentrations of cisapride in dogs (left panel) and human (right panels). The solid line represents the regression for the population predicted slope and intercept. Dots show the observed QTc-interval and the corresponding (predicted) individual concentration data.

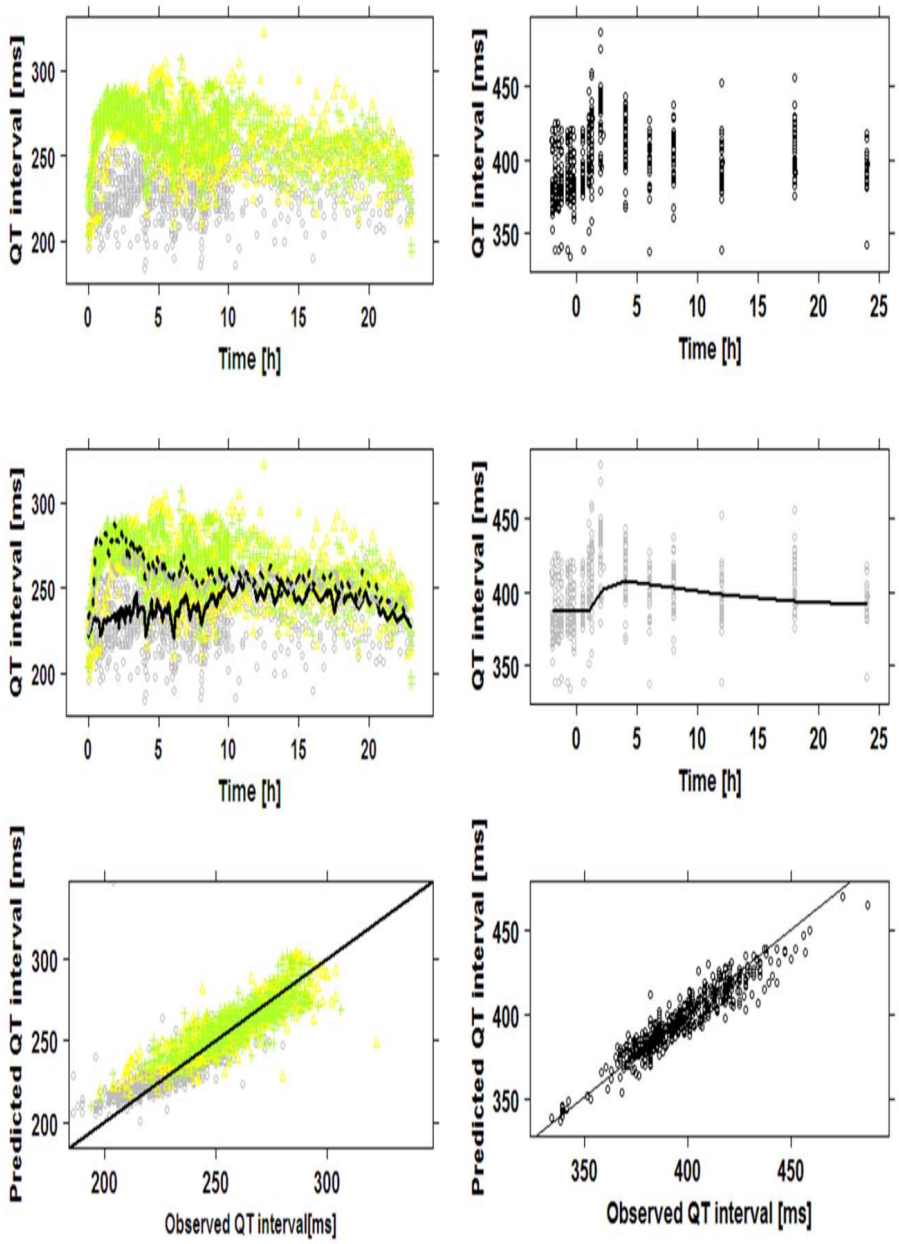


Figure 2b. Pharmacokinetic-pharmacodynamic relationship between QTc-interval and plasma concentrations of d,l- sotalol in dogs (left panel) and human (right panels). The solid line represents the regression for the population predicted slope and intercept. Dots show the observed QTc-interval and the corresponding (predicted) individual concentration data.

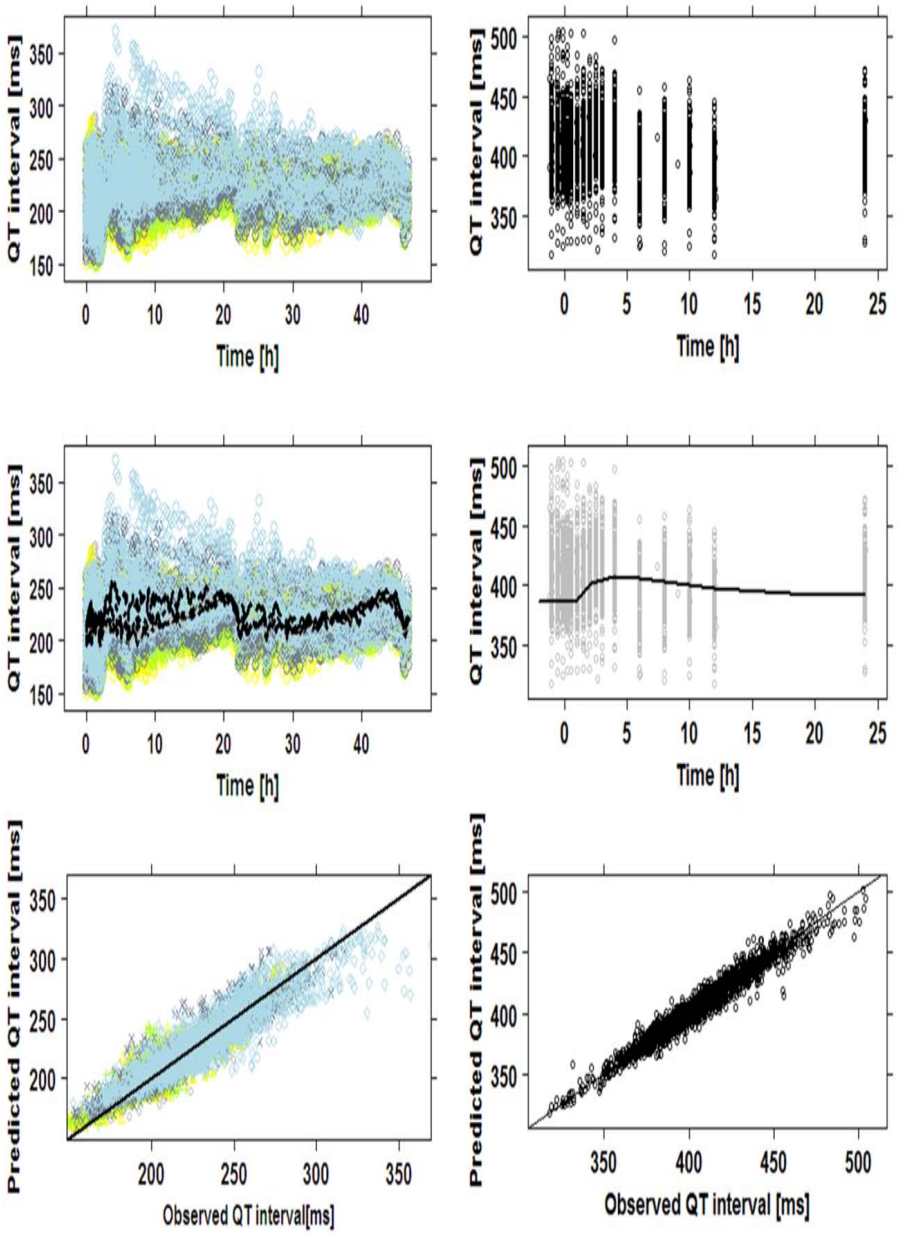


Figure 2c. Pharmacokinetic-pharmacodynamic relationship between QTc-interval and plasma concentrations of moxifloxacin in dogs (left panel) and human (right panels). The solid line represents the regression for the population predicted slope and intercept. Dots show the observed QTc-interval and the corresponding (predicted) individual concentration data.

Translational pharmacology

One of the main aspects our modelling exercise was to explore the predictive value of establishing PKPD relationships in dogs as basis for translating drug-induced QTc-interval prolongation from animals to humans. As shown in Figure 3, the use of a common model parameterisation allowed direct comparison of drug-induced effects across species. As depicted by the goodness-of-fit plots showing individual observed vs. predicted QTc-intervals, PKPD modelling showed similar performance in dogs and humans. Using Bayesian inference, it was possible to estimate the cumulative distribution of posterior samples for the differences between QTc baseline and QT after drug administration. The probability of QTc prolongation greater or equal to 10 msec was then calculated for each compound. Interestingly, the probability curve in humans clearly show a steeper increase across the therapeutic concentration range of each compound, indicating the risk of greater or equal to 10 msec QTc prolongation at such exposure levels. On the other hand, the probability of greater of equal to 10 msec QTc prolongation reaches maximum value consistently at higher exposure in dogs than in humans, suggesting potential differences in the sensitivity to the QT-prolonging effects across species.

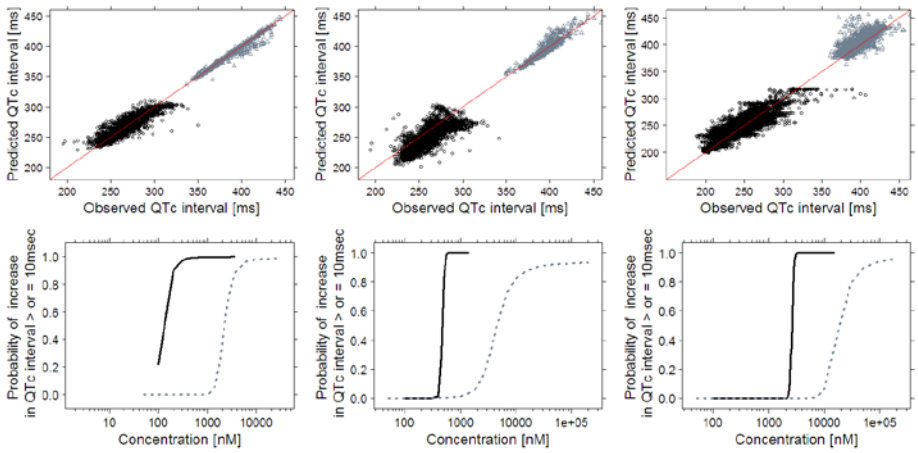


Figure 3. Top panels - Individual observed versus predicted QTc-intervals for cisapride (left), d,l-sotalol (middle) and moxifloxacin (right) for dogs and humans. Bottom panels - Corresponding probability curves for QTc-interval prolongation ≥ 10 msec versus predicted plasma concentrations for cisapride (left), d,l-sotalol (middle) and moxifloxacin (right). Black dots and dotted lines represent values for dogs while grey dots and solid lines are for human.

Significant differences were found in the slope parameters between the preclinical studies and clinical trials for all 3 compounds. The Z-scores were 3.06, 4.47 and 8.26 for cisapride, *d,l*-sotalol and moxifloxacin respectively.

4.4 DISCUSSION

PKPD modelling

In this study we have shown how a Bayesian hierarchical model can be used 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 humans at comparable exposure ranges. Thanks to the explicit distinction between drug-specific and system-specific parameters, our findings suggest that a model-based analysis of pharmacokinetic and QT-interval data in dogs allows inferences to be made about the probability of drug-induced QTc prolongation in humans.

Although a direct comparison with published data is not possible, it seems that dogs are slightly less sensitive to the QT-prolonging effects, as indicated by the somewhat less steep slope of the concentration-effect curves and lower probability values observed at concentration levels relevant to human therapeutic use. It was also evident from the analysis, that accurate interpretation of the preclinical findings require some knowledge about the expected therapeutic range, which was known and well established for the three reference compounds. This is reflected by the estimates of the drug-specific parameter in this model (slope), which does not independently indicate whether there is a risk associated with the use of the compounds, nor whether their development should be stopped. Rather, it provides an unambiguous description of the relationship between concentration and QT-prolonging effects taking into account various sources of variability.

From a drug development perspective our approach appears to overcome one of the main limitations of earlier PKPD modeling 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. Previous modelling attempts have focused on parameter estimation and goodness-of-fit to describe the observed data , rather than on

the use of models as a tool for predictive purposes [26, 27]. Some recent examples of model-based analysis of QTc prolongation often do not correlate findings to preclinical effects in a direct and quantitative manner [26, 27]. As illustrated by Ollerstam et al., different PKPD models were required to describe preclinical and clinical data. In other cases, additional *in vitro* data was required to correlate findings across species [26]. Furthermore, the choice of parameterisation in these analysis does not allow for reuse of the model for a different compound or compound class [26, 27], defeating the main purpose of such models, i.e., their use as a screening tool in early drug development.

Based on the evidence obtained so far, it appears that one of the most important factor for accurate translation of preclinical findings 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 QTc-interval and its clinical implication. In addition, our investigation showed some critical shortcomings in current experimental protocols which prevent effective use of model-based approaches in early drug development. Further consideration about these aspects are discussed in the subsequent paragraphs.

Experimental design requirements

The comparative analysis of preclinical and clinical data has provided the basis for further evaluation of the experimental requirements for the assessment of drug-induced QTc-interval prolongation using a model-based approach. Clearly, the main shortcoming lies in the way preclinical safety experiments are designed, both in terms of dose range and sampling schemes. Furthermore, the sample size should be considered more carefully when extrapolating estimates from preclinical safety experiments. A traditional cardiovascular safety study in conscious dogs has a Latin square, four-way crossover design, using four animals only. In this study we used a larger number of animals to ensure accurate parameter estimation, including estimates of between-subject variability. Yet, variability in dogs was found to be much larger than what is commonly observed in clinical trials. Various procedural differences may explain such differences, including the fact that dogs are freely moving when performing the ECG scan, while subjects are resting in supine or semi-supine position in clinical studies [2].

With regard to the use of empirical sampling schemes, it should be noted that the choice for limited or sparse sampling (usually 3-8 samples) without clear understanding of the pharmacokinetic properties in the species of interest renders the assessment of PKPD relationships rather difficult and imprecise. Optimal design concepts could be used to prevent or overcome these issues [28, 29]. In conjunction with nonlinear mixed effects modeling techniques, optimal sampling can be implemented to eliminate some of the technical challenges due to serial or frequent sampling (e.g. interference with PD measurements, total volume of blood per animal). Sampling matrices can also be considered in which some blood samples are collected during ECG monitoring and additional samples are taken at a separate occasion after or before the actual safety experiment. Other important confounding factors for the analysis of PKPD relationships in dogs are variation in age, body temperature and body weight of the animals, which are not collected with the same accuracy as comparable covariates in humans. Younger dogs have higher heart rates, thereby increasing overall noise:signal ratio and less precise baseline (intercept) QTc estimates [30].

Translational modelling of QTc-interval prolongation.

In drug development, the ideal situation of the prediction of drug-induced QTc prolongation would be if *in vitro* experiments would have a strong predictive value of the clinical outcome, both in qualitative (QTc prolongator or not) and quantitative (the amount of prolongation per concentration unit) terms. The value of a qualitative correlation (trying to avoid false positive and false negative predictions of QTc prolonging compounds) has been demonstrated by De Ponti et al. and Redfern et al., [31, 32]. One of the main issues with such qualitative approaches is that the potential small QTc prolonging compounds are hard to place. Furthermore, the QTc prolongation per concentration unit reached in humans cannot be predicted, making it impossible to establish recommendations for dosing regimen that yield a 'safe' QTc prolongation of less then 10 msec.

Although our work has focused on the translation of *in vivo* data, the availability of predictive measures which take *in vitro* properties into account remain 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.

We anticipate the possibility of further expanding the concept of a general model framework in which *in vitro-in vivo* correlations are assessed for different classses of compounds in terms of their ion-channel binding properties and used as 'scaling factors or covariates on drug-specific parameters (i.e. the slope). In fact, we have undertaken an initial evaluation of such correlations by analysing hERG assay data for the same compounds (data not shown). However, hERG assay was found to yield extremely variable results, making the use of *in vitro* data unreliable without a standardised protocol. As shown by Milnes et al., differences in experimental protocols have led to IC50 estimates for cisapride between 7mM and 72nM [33]. Additional complexities can also be anticipated for drugs showing multiple ion channel blocking properties, such as cisapride [31, 34]. Another issue is that the *in vitro* experiments using hERG essays only, would not detect the effects of Ca^{2+} channel blocking compounds. Martin et al. showed the masking effect of the Ca^{2+} channel inhibition on the APD while there was a large hERG inhibition measured, giving a QTc shortening at higher concentrations after the QT prolongation at lower concentrations [35].

An important aspect of our analysis was the possibility to better understand the mismatch often observed between preclinical and clinical findings, which is referred to a false-positive or false-negative results. Our findings suggest that the assessment of QTc-interval prolongation using different experimental protocols in preclinical species appear to be meaningless without clear understanding of the underlying exposure and the corresponding therapeutic range. On the other hand, thanks to the use of PKPD relationships as a common denominator across species, it is possible to explain in a more systematic manner whether intrinsic differences exist between dogs and humans and how such differences may be related to known physiological factors. For instance, it is conceivable that the differences in I_{kr} channel distribution in the left ventricular wall of the canine heart [36] may lead to differences in repolarisation of the heart cells and consequently to a different degree of QT prolongation [37]. The other factor is the difference in basal heart rate, which is considerably higher in dogs. To detect small differences in QT-interval, heart rate correction based on population-based correction factors, such as Bazett and Fridericia, may not be equally applicable to dogs [38, 39].

Limitations of the approach

Despite the clear correlation between dogs and humans as well as the goodness-of- fit, a potential limitation of our approach is the fact that it relies on a linear regression for the assessment of drug induced QTc prolongation. The estimation of a slope contrasts with the more commonly used sigmoidal concentration-effect relationship. In clinical trials the Emax (the maximum QTc prolongation) will rarely be observed as arrhythmias because the subjects are already pulled from the trials due to ECG stopping criteria being reached (e.g. QTc > 500 msec). When the Emax cannot be estimated in a proper fashion, a small misspecification of this value will have a huge effect on the EC50. As the difference of only 10 msec on a 400 msec average QT-interval is needed to increase the risk on TdP, the error on the estimation of EC50 could imply false positive or false negative results. The use of this linear concentration-effect relationship will thereby be more accurate in the lower effect range.

As highlighted previously, another limitation to this method is the requirement for higher quality data. When low quality data is used or sampling schemes are applied in an empirical manner, without taking into account drug absorption, distribution and elimination processes, the PK model misspecification can lead to to imprecision and consequently resulting in very wide confidence intervals for the slope parameter as derived in Equation 1.

In summary, the possibility of extrapolating data on QTc-interval prolongation across species and using the concentration-effect relationships in terms of the risk of QTc-interval prolongation $\geq$ 10 msec can greatly enhance the decision-making process for both the regulators and drug developers. In addition, our findings also show that accurate evaluation of the proarrhythmic potential of a novel chemical or biological entity cannot be performed without distinguishing drug-specific parameters from system-specific parameters. In this sense, we envisage the use of this Bayesian approach as the basis for translational pharmacology research, particularly at the interface between drug discovery and early drug development, making it possible to assess the clinical relevance of drug-induced effects in preclinical species.

REFERENCES

[1] Darpo B, Nebout T, Sager PT. Clinical evaluation of QT/QTc prolongation and proarrhythmic potential for nonantiarrhythmic drugs: the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use E14 guideline. J Clin Pharmacol 2006 May;46(5):498-507.

[2] Guidance for Industry: ICH E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrythmic Potential for Non-Antiarrythmic Drugs.U.S. Department of Health and Human Services. Food and Drug Administration. (2005)

[3] Danhof M, de Lange EC, Della Pasqua OE, Ploeger BA, Voskuyl RA. Mechanism-based pharmacokinetic-pharmacodynamic (PK-PD) modeling in translational drug research. Trends Pharmacol Sci 2008 Apr;29(4):186-91.

[4] Guidance for Industry: ICH S7A Safety pharmacology studies for human pharmaceuticals. U.S. Department of Health and Human Services. Food and Drug Administration. (2001)

[5] Guidance for Industry: ICH S7B The nonclinical evaluation of the potential for delayed ventricular repolarization (QT interval prolongation)by human pharmaceuticals. U.S. Department of Health and Human Services. Food and Drug administration. (2005)

[6] Fleury A, Lave T, Jonsson F, Schmitt M, Hirkaler G, Polonchuk L, et al. A pharmacokinetic-pharmacodynamic model for cardiovascular safety assessment of R1551. J Pharmacol Toxicol Methods 2011 Jan-Feb;63(1):123-33.

[7] Zareba W. Drug induced QT prolongation. Cardiol J 2007;14(6):523-33.

[8] Roden DM. Drug-induced prolongation of the QT interval. N Engl J Med 2004 Mar 4;350(10):1013-22.

[9] Cavero I, Mestre M, Guillon JM, Crumb W. Drugs that prolong QT interval as an unwanted effect: assessing their likelihood of inducing hazardous cardiac dysrhythmias. Expert Opin Pharmacother 2000 Jul;1(5):947-73.

[10] Hammond TG, Carlsson L, Davis AS, Lynch WG, MacKenzie I, Redfern WS, et al. Methods of collecting and evaluating non-clinical cardiac electrophysiology data in the pharmaceutical industry: results of an international survey. Cardiovasc Res 2001 Mar;49(4):741-50.

[11] Netzer R, Ebneth A, Bischoff U, Pongs O. Screening lead compounds for QT interval prolongation. Drug Discov Today 2001 Jan 1;6(2):78-84.

[12] Malik M, Andreas JO, Hnatkova K, Hoeckendorff J, Cawello W, Middle M, et al. Thorough QT/QTc study in patients with advanced Parkinson's disease: cardiac safety of rotigotine. Clin Pharmacol Ther 2008 Nov;84(5):595-603.

[13] Piotrovsky V. Pharmacokinetic-pharmacodynamic modeling in the data analysis and interpretation of drug-induced QT/QTc prolongation. AAPS J 2005;7(3):E609-24.

[14] Vandemeulebroecke M, Lembcke J, Wiesinger H, Sittner W, Lindemann S. Assessment of QT(c)-prolonging potential of BX471 in healthy volunteers. A 'thorough QT study' following ICH E14 using various QT(c) correction methods. Br J Clin Pharmacol 2009 Sep;68(3):435-46.

[15] Chain AS, Krudys KM, Danhof M, Della Pasqua O. Assessing the probability of drug-induced QTc-interval prolongation during clinical drug development. Clin Pharmacol Ther 2011 Dec;90(6):867-75.

[16] Lunn DJ, Thomas, A., Best, N & Spiegelhalter, D. WinBUGS - A Bayesian modelling framework: concepts, structure, and extensibility. Stat Comput 2000;10:325-37.

[17] Spiegelhalter D, Thomas, A., Best, N. & Gils, W. BUGS: Bayesian Inference Using Gibbs Sampling, Version 1.4.1. Cambridge, UK: MRC Biostatistics Unit; 2003.

[18] Lunn DJ, Best N, Thomas A, Wakefield J, Spiegelhalter D. Bayesian analysis of population PK/PD models: general concepts and software. J Pharmacokinet Pharmacodyn 2002 Jun;29(3):271-307.

[19] Duffull SB, Kirkpatrick, C.M., Green, B. & Holford, N.H. Analysis of population pharmacokinetic data using NONMEM and WinBUGS. J Biopharm Stat Comput 2005;15:53-73.

[20] Bachman DM. Oral magnesium ion shortens prolonged QTc interval. The Journal of clinical psychiatry 2003 Jun;64(6):733-4.

[21] Spiegelhalter D AK, Myles J. . Bayesian approaches to clinical trials and health care evaluations New York,: Wiley; 2004.

[22] Spiegelhalter D, Best, N., Carlin, B. . Bayesian measures of model complexity and fit. . J Roy Stat Soc 2002;Ser B(583-639).

[23] Brooks SPG, A. A General methods for monitoring convergence of iterative simulations. J COmput Graph Stat 1998:434-55.

[24] Gelman A, Carlin, J.B., Stern, H.S. & Rubin, D.B. Bayesian Data Analysis. London: Capman & Hall; 2003.

[25] Gelman AR, D. Inference from iterative simulation using multiple sequences. Stat Sci 1992:457-511.

[26] Jonker DM, Kenna LA, Leishman D, Wallis R, Milligan PA, Jonsson EN. A pharmacokinetic-pharmacodynamic model for the quantitative prediction of dofetilide clinical QT prolongation from human ether-a-go-go-related gene current inhibition data. Clin Pharmacol Ther 2005 Jun;77(6):572-82.

[27] Ollerstam A, Visser SA, Persson AH, Eklund G, Nilsson LB, Forsberg T, et al. Pharmacokinetic-pharmacodynamic modeling of drug-induced effect on the QT interval in conscious telemetered dogs. J Pharmacol Toxicol Methods 2006 Mar-Apr;53(2):174-83.

[28] Dodds MG, Hooker AC, Vicini P. Robust population pharmacokinetic experiment design. J Pharmacokinet Pharmacodyn 2005 Feb;32(1):33-64.

[29] Nyberg J, Karlsson MO, Hooker AC. Simultaneous optimal experimental design on dose and sample times. J Pharmacokinet Pharmacodyn 2009 Apr;36(2):125-45.

[30] Salvi V, Karnad DR, Panicker GK, Natekar M, Hingorani P, Kerkar V, et al. Comparison of 5 methods of QT interval measurements on electrocardiograms from a thorough QT/QTc study: effect on assay sensitivity and categorical outliers. J Electrocardiol 2011 Mar-Apr;44(2):96-104.

[31] Redfern WS, Carlsson L, Davis AS, Lynch WG, MacKenzie I, Palethorpe S, et al. Relationships between preclinical cardiac electrophysiology, clinical QT interval prolongation and torsade de pointes for a broad range of drugs: evidence for a provisional safety margin in drug development. Cardiovasc Res 2003 Apr 1;58(1):32-45.

[32] De Ponti F, Poluzzi E, Montanaro N. Organising evidence on QT prolongation and occurrence of Torsades de Pointes with non-antiarrhythmic drugs: a call for consensus. Eur J Clin Pharmacol 2001 Jun;57(3):185-209.

[33] Milnes JT, Witchel HJ, Leaney JL, Leishman DJ, Hancox JC. Investigating dynamic protocol-dependence of hERG potassium channel inhibition at 37 degrees C: Cisapride versus dofetilide. Journal of pharmacological and toxicological methods 2010 Mar-Apr;61(2):178-91.

[34] Davie C, Pierre-Valentin J, Pollard C, Standen N, Mitcheson J, Alexander P, et al. Comparative pharmacology of guinea pig cardiac myocyte and cloned hERG (I(Kr)) channel. Journal of cardiovascular electrophysiology 2004 Nov;15(11):1302-9.

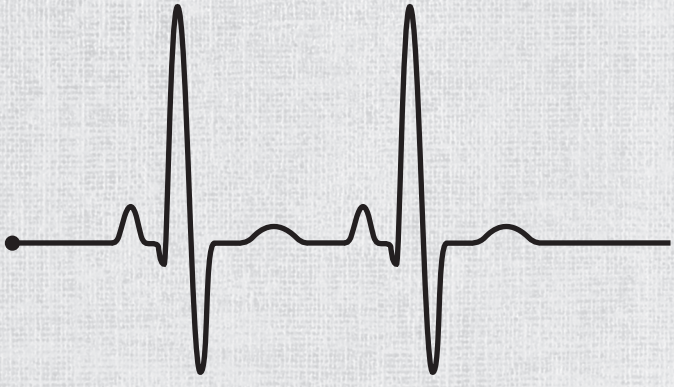
[35] Martin RL, McDermott JS, Salmen HJ, Palmatier J, Cox BF, Gintant GA. The utility of hERG and repolarization assays in evaluating delayed cardiac repolarization: influence of multi-channel block. J Cardiovasc Pharmacol 2004 Mar;43(3):369-79.

[36] Gintant GA. Regional differences in IK density in canine left ventricle: role of IK,s in electrical heterogeneity. The American journal of physiology 1995 Feb;268(2 Pt 2):H604-13.

[37] Antzelevitch C, Sicouri S, Litovsky SH, Lukas A, Krishnan SC, Di Diego JM, et al. Heterogeneity within the ventricular wall. Electrophysiology and pharmacology of epicardial, endocardial, and M cells. Circulation research 1991 Dec;69(6):1427-49.

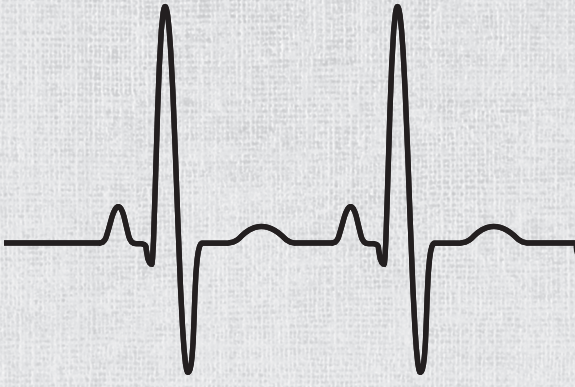
[38] Soloview MV, Hamlin RL, Barrett RM, Chengelis CP, Schaefer GJ. Different species require different correction factor for the QT interval. Cardiovascu Toxicol 2006; 6(1):51-52.

[39] Markert M, Shen R, Trautmann T, Guth B. Heart rate correction models to detect QT interval prolongation in novel pharmaceutical development. J Pharmacol Toxicol Methods 2011 Jul-Aug;64(1):25-41.



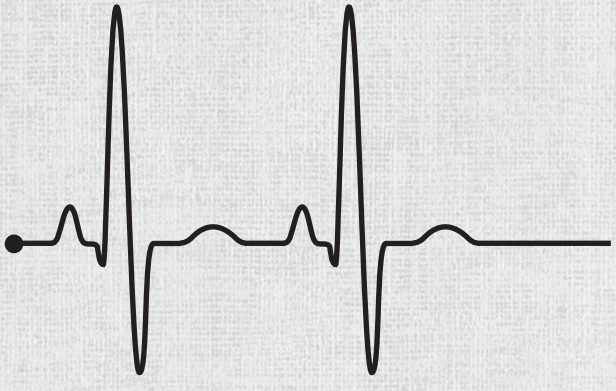
"True wisdom is knowing what you don't know."

Confucius



SECTION III

FROM PHARMACOLOGY TO EPIDEMIOLOGY - A BRIDGING STRATEGY FOR RISK MANAGEMENT

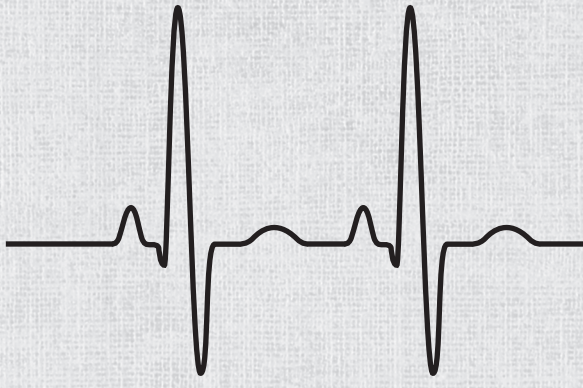


Not-In-Trial Simulation I:
Bridging Cardiovascular Risk From Clinical Trials To Real Life Conditions

ASY Chain (1, 2), JP Dieleman (2), C van Noord (3), M Danhof (1), MCJM Sturkenboom (2, 3), O Della Pasqua (1,4)

(1) Leiden/Amsterdam Centre for Drug Research, Division of Pharmacology, Leiden University, Leiden, The Netherlands
(2) Department of Medical Informatics and Epidemiology, Erasmus Medical Centre, The Netherlands
(3) Department of Epidemiology, Erasmus Medical Centre, The Netherlands
(4) Clinical Pharmacology and Discovery Medicine, GlaxoSmithKline, Stockley Park, Uxbridge, United Kingdom

British Journal of Clinical Pharmacolgy - in press



Chapter 5

ABSTRACT

Given the assumed risk of proarrhythmias, assessment of QTc-interval prolongation has become a mandatory step in drug development. Currently, however, experimental data are generated from randomised trials in healthy volunteers. The present study was performed to explore discrepancies between drug-induced QTc prolongation in clinical trials and real life conditions. d,l-Sotalol was selected as a paradigm compound to illustrate the potential clinical implications of such differences. The Rotterdam Study cohort was used as the real life reference population. Using clinical trial simulation principles and nonlinear mixed effects modelling, we show that significant differences exist between the drug-induced increase in QTc-interval and observed QTc values. Furthermore, our analysis reveals that the increased prevalence of high QTc values in a real-life population is also due to co-morbidities and concomitant medications, rather than treatment drug alone. These findings substantiate the need to account for these factors when evaluating cardiovascular risk of novel medicinal products.

5.1 INTRODUCTION

In the last two decades, there have been increased concerns around the rising number of proarrhythmia cases due to non-antiarrhythmic drug-induced QTc prolongation to the extent that it has been referred to as a modern "epidemic" [1]. Physiologically, the duration of the QT-interval of the surface electrocardiogram (ECG) represents the ventricular action potential duration. As such, prolonged heart-rate corrected QT- (QTc-) interval may result in early after-depolarisations (EAD), which may induce re-entry, provoke torsade de pointes (TdP) and potentially lead to sudden cardiac death (SCD) [2, 3].

Even though epidemiological evidence regarding the correlation between drug-induced QTc-interval prolongation, risk of fatal arrhythmias and SCD remains under scrutiny [4, 5], QTc prolongation has become the second most common cause for post-market drug withdrawal [6]. In other cases, severe prescribing restrictions have been imposed by regulatory agencies, which impacted on the indication of the drug, its posology, contraindications, special warnings and precautions for use, such as baseline and periodic on-treatment ECGs monitoring. The requirements for information on drug-disease and drug-drug interactions, full description of cardiotoxic adverse effects and their management as well as risk of overdose have also been affected [7]. Not only have the regulatory bodies started to closely monitor the liability for QTc-interval prolongation in the submissions of new chemical entities (NCE), guidelines have also emerged outlining the pre-requisites for the evaluation of drug-induced effects on heart conductivity and QTc-interval during drug development.

The first regulatory guidelines regarding clinical evaluation of the QT/QTc-interval prolongation in the context of new drug development were issued in 1997 by the Committee for Proprietary Medicinal Products (CPMP) of the European Union [8, 9]. More recently, the International Conference on Harmonisation (ICH) Topic E14 guideline was published on the clinical evaluation of QT/QTc-interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs [10]. The ICH guideline introduced the use of the "double-delta" method in a thorough-QT study (TQT) as the state-of-the-art approach for the assessment of QTc-interval prolongation. Furthermore, it recommends a TQT to be completed early in the drug development process to guide further studies and decision making.

From a methodological perspective, the TQT protocol design relies on the use of a positive control as well as supra-therapeutic doses of the investigational drug. The outcome of the analysis is based on the upper bound of the 95% one-sided confidence interval for the largest time-matched mean effect of the drug on the QTc-interval, which may not exceed 10 msec for the study to be deemed negative. Whilst the validity as well as the sensitivity of the proposed methodology has been the subject of debate in numerous publications, less attention has been paid to other aspects of the preferred protocol design and interpretation of the findings, which overlook the underlying concentration-QTc effect relationship and consequently supporting evidence of causality [11, 12].

Taking into account concentration-effect relationships, we explore the implications of the preferred protocol design for the assessment of cardiovascular risk under real life conditions. Although it may be understandable and desirable to characterise a drug safety profile in healthy subjects, we show how population selection and inclusion/exclusion criteria lead to a somewhat biased view of drug-induced effects in the target patient population which will receive the treatment after approval. The selected trial population, strict inclusion/exclusion criteria and experimental conditions may make the extrapolation of the findings to a real life population rather cumbersome. In particular, the exclusion of severely ill subgroups, i.e., those who are most susceptible to the potential undesired proarrhythmic effect, is usually not factored in a quantitative manner when evaluating TQT results. In addition, little attention is paid to the role of other factors contributing to QTc-interval prolongation in the target population, as compared to the observed drug effect in a randomised clinical trial involving healthy subjects. The current approach to clinical trials is also too focused on designs aimed at demonstrating efficacy. The number of subjects exposed to the NCE is powered to show benefit rather than to pick up signals from rare but potentially fatal adverse events and no formal procedures exist to mitigate the impact of such differences or support the management of cardiovascular risk in the target population. Finally, many patient subgroups with higher risk factors for TdP are often excluded from the trials. These subgroups include females, elderly subjects, those with predisposed cardiac or non-cardiac diseases associated with diminished repolarisation reserve (and therefore greater susceptibility to QT-interval prolongation), those with pharmacogenetic defects of drug metabolising enzymes or pharmacological targets such as the potassium channels, those susceptible to bradycardia or

electrolyte imbalance, or those receiving drugs with a potential for pharmacokinetic or pharmacodynamic interactions [13, 14].

From a regulatory perspective, questions have also arisen about the efficiency and reliability of pre-approval clinical trials in identifying the clinical risk of torsade de pointes [12, 15]. Given the patient population enrolled, the background noise (arising from spontaneous intra-individual variability in QTc-interval) and the relatively low frequency of clinically significant drug–induced effects, clinical trials may or may not accurately detect the frequency and intensity of QTc-interval prolongation. In fact, it is known that the pro-arrhythmic threshold can vary across compounds with frequency of such events ranging from approximately 1 in 100 (for halofantrine) to 1 in 50000 (for terfenadine) [1].

The evolving concepts in risk management will inevitably lead sponsors, regulatory agencies and other stakeholders to consider how to best evaluate causality and identify the contribution of other factors determining increases in QTc-interval and consequently in cardiovascular risk in the target population. In this investigation we characterise QTc-interval changes following administration of *d,l*-sotalol (paradigm compound) to healthy subjects in clinical trials and to patients in real life. Our investigation resorts to model-based drug development (MBDD) principles, as the development and application of pharmaco-statistical models of drug efficacy and safety from preclinical and clinical data, to improve drug development knowledge management and decision making [16, 17]. In contrast to typical clinical trial simulations [18-20], we make use of simulations to characterise the role of design factors which have been omitted or excluded from a randomised trial, among which those listed under the exclusion criteria or diverging from the clinical conditions relevant for the target population. Thus, we anticipate that our approach will represent a natural extension of ongoing efforts within pharmaceutical industry to improve safety signal detection [21-25].

5.2 METHODS

Real life study population

The real life population setting was derived from the Rotterdam Study. It is a prospective population-based cohort study, which started with a baseline

visit between 1990 and 1993. All inhabitants of a suburb in Rotterdam, Ommoord, aged 55 years and over were invited to participate. Of the 10 275 subjects invited, 7983 (78%) gave their written informed consent and took part in the baseline examination. The study was approved by the Medical Ethics Committee of the Erasmus Medical Center, Rotterdam, The Netherlands. Objectives and methods of investigation have been described in detail in other publications [35]. All participants were visited at home for a standardised questionnaire and 7151 were subsequently examined at the research center. A second, third and fourth follow up visit took place, respectively, between 1993 - 1995, 1997 - 1999 and 2002 - 2004. In addition to follow-up examinations, the cohort is continuously monitored for major morbidity and mortality through linkage with general practitioner and municipality records. Drug prescriptions dispensed to participants by automated pharmacies have been routinely stored in the database since 1 January, 1991. For the present study, we included all participants in the Rotterdam Study with at least two consecutive ECG assessments.

All subjects who started treatment with *d,l*-sotalol during follow-up were included in the this study, with the exception of prevalent users and subjects with left ventricular hypertrophy (LVH), right bundle-branch block (RBBB) and left bundle-branch block (LBBB), who were excluded. The reason for the exclusion of these patients is that LVH, RBBB and LBBB can cause secondary repolarisation changes and atrial fibrillation can cause difficulties in measuring QT-intervals. From the remaining population, "new" *d,l*-sotalol users were identified as those receiving their first prescription after enrolment into the study. Our analysis included all QTc measurements after the start of *d,l*-sotalol for all new users.

During the research center visit, non-fasting blood samples were obtained [36]. Body mass index (BMI) was computed as weight divided by height squared. Comorbidities were screened and identified according to standard clinical criteria. Hypertension was defined as systolic blood pressure > 160 mm Hg or diastolic blood pressure > 100 mm Hg and/or use of antihypertensive medications, encompassing grade 2 and grade 3 hypertension compounds according to the World Health Organisation (WHO) criteria [37]. Diabetes mellitus was defined as the use of blood glucose-lowering medication and / or a non-fasting or post-load serum glucose level of 11.1 nmol/L or higher according to the WHO [38]. A history of myocardial infarction was assessed by self-report checked with records from the general

practitioner or cardiologist and/or electrocardiographic evidence. All reported myocardial infarctions were verified against medical records as described in detail previously [39]. Assessment of heart failure at baseline and during follow-up was assessed by reviewing all medical records for the occurrence of at least two signs and symptoms suggestive of heart failure or the use of medication for the indication heart failure and hospital discharge letters. Cases of incident heart failure were obtained by continuously monitoring the participants [40, 41]. The ankle arm index (AAI) was used as potential predictor of cardiovascular diseases and mortality [42].

Healthy volunteer study population

A healthy volunteer clinical trial for *d,l*-sotalol was selected from GlaxoSmithKline's clinical data repository (study number EXP20001). The primary objective of this study was to assess the role of intrinsic and extrinsic factors on the variability of QTc-interval and was powered to detect a 10 msec increase in QTc according to the ICH E14 double-delta method. The study had a placebo-controlled, three-way crossover design with two placebo periods and one active treatment period. Blood concentrations were taken at -30, 5, 15, 30 and 45 minutes relative to dosing time as well as 1, 2, 4, 8, 10, 18 and 24 hours after dose. The study population comprised of 30 subjects (12 females, 18 males), between the ages of 19 and 47.

Relevant eligibility criteria for inclusion into (items 1-3) / exclusion from (items 4-8) the study for the purposes of our investigation included:

1. Age between 18-55 years, inclusive

2. Non-smoker (subject must have been a non-smoker for at least 3 months prior to screening)

3. Body mass index between 19-30 kg/m^2, with a weight of 50 – 95 kg, inclusive

4. At screening ECG recording showed abnormal QRST complex morphology, sinus bradycardia (HR<45 bmp, PR > 210 msec) or QTc-interval values above 420 or 440 msec for male and female, respectively.

5. Any medical history which was contra-indicated in the SOTACOR™ product label

6. History of hypertension, asthma, bronchial hyperactivity or peripheral vascular disease, including diabetes.

7. The subject had a history of alcohol or drug abuse

8. The subject was positive for HIV, hepatitis C or hepatitis B surface antigen

Drug-induced QTc-interval prolongation

The primary objective of this study was to identify differences in the baseline QTc-interval (i.e. before *d,l*-sotalol use) between target patient population and healthy volunteers. The simulation of QTc-intervals was performed in a two-step approach. First, drug concentrations in the target population (i.e., *d,l*-sotalol users) were derived using a pharmacokinetic model. QTc-intervals were subsequently obtained using final population pharmacokinetic-pharmacodynamic (PKPD) parameter estimates.

The pharmacokinetics of *d,l*-sotalol was described by a two compartment, first-order absorption model with weight as a covariate on clearance using nonlinear mixed effects modelling, as implemented in NONMEM (v5.1, ICON Development Solutions**,** Ellicott City, MD, USA) [43]. Heart-rate corrected QT-intervals were described by a PKPD model using a Bayesian approach in which physiological (QT and RR) and drug specific (concentration) parameters are parameterised in an independent manner (Equation 1) [44]. A detailed description of model development was described in a previous publication [11].

$$QT_c = QT_0 \cdot RR^{\alpha} + A \cdot \cos\left(\frac{2\pi}{24}(t - \phi)\right) + slope \cdot C$$

Equation 1

where, QT_0 [msec] is the intercept of the QT-RR relationship. Sex was included as a covariate for this parameter whenever applicable. RR [sec] is the interval between successive R waves, α is the individual heart rate

correction factor, A [msec] is the amplitude of circadian rhythm, t is the clock time, φ is the phase, slope [msec/concentration] is the linear pharmacodynamic relationship, and C is the predicted concentration of drug at the time of QT-interval measurements.

Drug-induced QTc-intervals in the Rotterdam Study population were then simulated using the PKPD model, taking into account patient characteristics in the real life population. The QTc-interval distribution associated with *d,l*-sotalol exposure in the reference population was simulated assuming full compliance to treatment. Population parameter estimates were used for the simulations (slope = 0.02 msec/ng/ml, intercept (QT_0) = 380 msec). Finally, QTc values in the Rotterdam Study population were compared non-parametrically with the observed data.

ECG measurements

For the Rotterdam Study cohort, a 10 sec 12-lead ECG, resulting on average of 8-10 heart beats, was recorded with an ACTA electrocardiograph (ESAOTE, Florence, Italy) at a sampling frequency of 500 Hz and stored digitally. All ECGs were processed by the Modular ECG Analysis System (MEANS) to obtain the ECG parameters of interest. The MEANS program has been evaluated extensively and described elsewhere [45, 46]. MEANS determines common onset and offset for all 12 leads together for one representative averaged beat, with the use of template matching techniques, until the end of the T wave. To adjust for heart rate, Bazett's formula (QTc = QT / RR ^ ½) was used [47]. A total of 1387 digitally stored ECG records were used for this study. On the other hand, ECG measurements in healthy volunteers were recorded in triplicate and read manually from lead II by a blinded cardiographist. The average value from the triplicates was used for data analysis and development of the PKPD model.

Statistical analysis

Differences between clinical trials and real-life population:

Observed and expected QTc-intervals were non-parametrically compared visually and statistically, by determining the significance of the differences between the two distributions. Specifically, the Wilcoxon signed-rank test was used to evaluate if the distributions significantly differ from each other with p-value < 0.05.

Covariate effects:

The relative risk (RR) and 95% confidence interval (CI) of the association between "prolonged" QTc, as defined as > 450 ms for males and > 470 msec for females, and concomitant use of medications and presence of comorbidities while on *d,l*-sotalol prescription were estimated using Cox's proportional hazards analysis [48]. At each ECG measurement, all information concerning medications and comorbidities was updated. For the purpose of this analysis, each measurement from the same subject was treated independently since within-subject variability could not be considered in this analysis.

5.3 RESULTS

After applying the target population selection criteria to the Rotterdam cohort as described in the *methods* section, 608 subjects were classified as "new" *d,l*-sotalol users without left ventricular hypertrophy (LVH), right bundle-branch block (RBBB) or left-bundle-branch block (LBBB). The total accumulated follow-up was 4864 years (i.e., an average and standard deviation of 8 $\pm$ 1.5 years/subject), during which they all had at least one ECG measurement. At the end, there were a total of 1375 available ECGs, with an average of two per person. Inclusion and exclusion criteria from the phase I *d,l*-sotalol trials were then applied to this subset to allow further characterisation of the differences between the real life conditions and the clinical trial protocols. Figure 1 reveals that 10.8% of the male population and 3.4% of the female population from the Rotterdam cohort, would have been excluded based on their weight values, which is a covariate used in the clinical pharmacokinetic-pharmacodynamic (PK/PD) model. Similarly, 21.9% male and 14.9% female would not be included due to their baseline QT-interval measurements (Figure 2).

Pharmacokinetic and PKPD model parameters were used to simulate *d,l*-sotalol concentrations and corresponding QT-intervals and to test if drug effects alone can explain the observed QTc values in observational real life cohort. Drug concentrations were simulated taking into account the individual weights of the elderly population. The simulated QTc-intervals (i.e. the expected QTc) were found to be at variance with the observed values (Figure 3), especially in the male population. Using Wilcoxon T-test, the

results showed that the two distributions significantly differed from each other for both genders with p-values < 0.05. The distribution of QTc values of the *d,l*-sotalol users in the real life elderly population revealed some dangerously high observed measurements. The model-based predictions for a real life population demonstrated that drug-induced effects alone cannot describe the observed values in the QTc-interval distribution, especially in men.

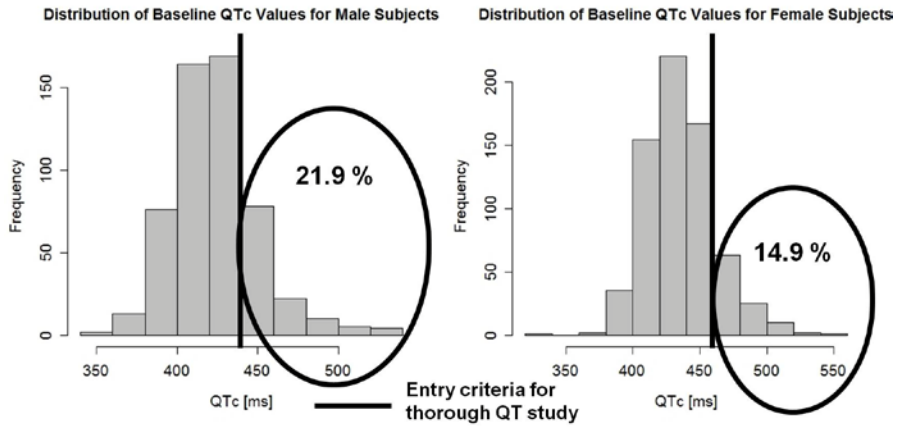


Figure 1. Weight distribution of males and females in the Rotterdam Study cohortt.

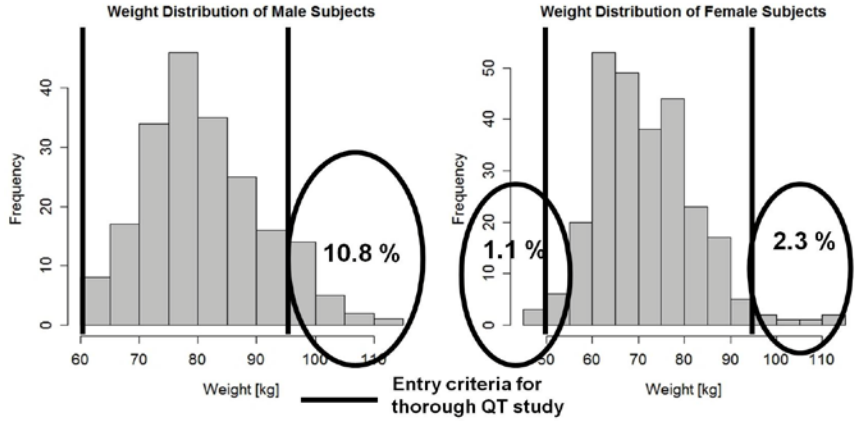


Figure 2. Baseline distribution of males and females in the Rotterdam Study cohort.

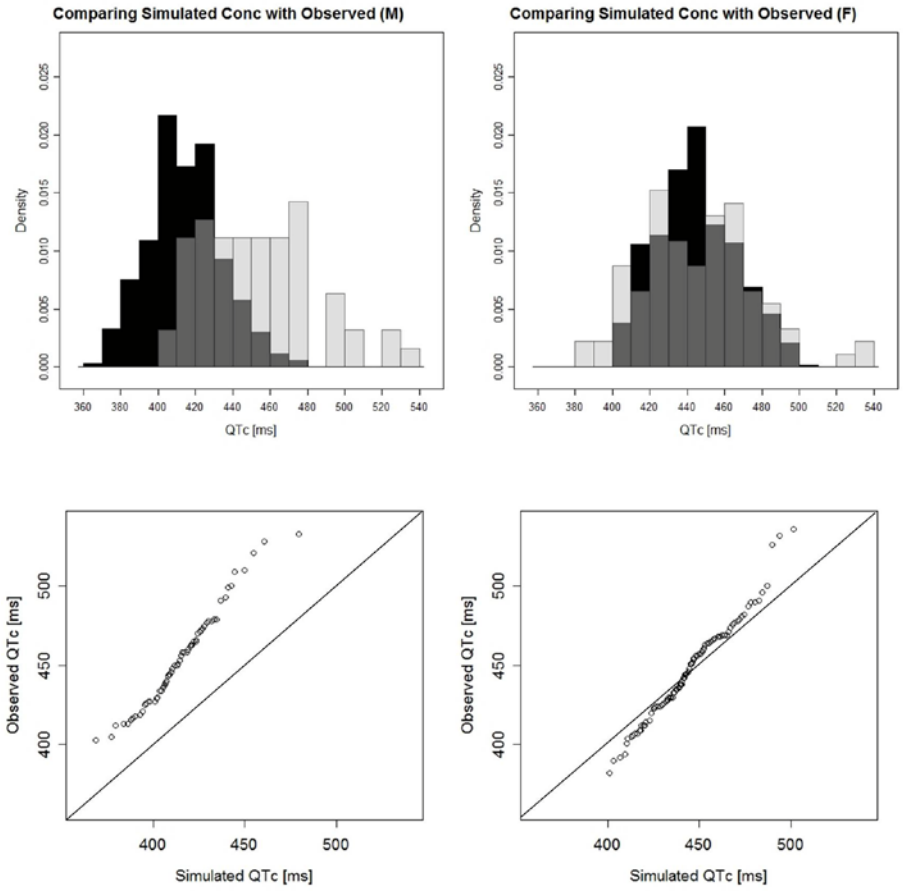


Figure 3. (a) Overlapping distributions showing the discrepancies between observed and simulated results in the male population (left panel) and female population (right panel). The darker colour represents the simulated QTc values and the light colour represents the observed QTc-intervals. The medium shades denote the overlapping areas. (b) QQ-plots comparing the distributions of the QTc-values for male population (left) and female population (right). The deviation from the identity line reflects the residual difference between the observed QTc-intervals and model-predicted d,l-sotalol effects under the assumption of comparable pharmacokinetic-pharmacodynamic relationship, as determined in Phase I clinical trials.

The differences between the two populations were evaluated by binary logistic regression, which revealed the contribution of concomitant medications and comorbidities as QTc-prolonging factors. It was found that heart failure, diabetes mellitus, myocardial infarction and hypertension increase the risks of a male having QTc >450 msec and female with QTc >470 msec by 0.4-11.5 fold (Table 1).

Comed/ Comorbidity	RR without condition (95% CI)	p-value	RR with condition (95% CI)	p-value
Using *d,l-sotalol* alone	-	< 0.01	4.0 (2.7 - 5.8)	< 0.01
Diabetes	4.0 (2.7 - 5.8)	< 0.01	6.5 (1.6 - 27.1)	< 0.01
Heart Failure	4.4 (3.0 - 6.6)	< 0.01	0.8 (0.2 - 3.4)	> 0.01
Hypertension	7.4 (4.3 - 12.7)	< 0.01	2.4 (1.3 - 4.2)	< 0.01
Myocardial Infarction	3.4 (2.3 - 5.127)	< 0.01	15.5 (4.9 - 49.3)	< 0.01

Table 1. Results of relative risk calculations of the association between "prolonged" QTc and concomitant use of medications and presence of comorbidities while under d,l-sotalol prescription. Significance level was set at $p < 0.05$.

5.4 DISCUSSION

In the current investigation 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. On the other hand, the signal detected during a TQT study has been deemed sensitive enough to reflect the increase in cardiovascular risk associated with torsadogenic or arrythmogenic effects of a drug in the target patient population.

Thanks to the use of a model-based approach it was possible to make inferences about drug exposure in patients and evaluate in an integrated manner, how different covariates and sources of variability affect the observed QTc values in real life patients. Our findings suggest that the distribution of observed QTc values in the real life cohort cannot be explained by the effects of *d,l*-sotalol alone. Other causal factors are present, which significantly affect the observed QTc values. In fact, previous publications showed that heart failure, hypertension, diabetes and myocardial infarction increase the risk of QTc prolongation [26-29]. The binary logistic regressions performed on our data revealed that diabetes, heart failure, 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 co-medication, both of which contribute to upper tail of the distribution of observed QTc values.

The concepts outlined in the current investigation also show how both clinical trial and epidemiological data can be used in an integrated manner for the purpose of signal detection and improved risk management. Our findings appear to confirm the points highlighted by Black on "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." [30] Others have also advocated the synergistic potential for using both kinds of data to aid decision making [31-34]. In the design of this study, we apply a parametric approach in which simulations play a central role. As a matter of fact, this is the first time that simulations based on nonlinear hierarchical models are used to characterise the implications of exclusion criteria on overall treatment safety profile.

By applying PKPD modelling concepts to epidemiological data, we have also shown that the concentration-effect relationship achieved from a healthy volunteer trial cannot adequately describe nor predict what is actually happening in real life. The results from the present study suggest that new methodology is required to assess the impact of QTc prolongation in real life population for future compounds. By extension, the relevance of a TQT study, as mandated in the ICH E14 guideline, becomes questionable. Our findings clearly challenge the validity of TQT trials as the state-of-the-art approach for determining the risk of pro-arrhythmia or TdP. Indirectly, our analysis also reveals the shortcomings of the double-delta method currently suggested for the TQT studies. The use of time-matched differences between active and placebo arms does not allow the assessment of potential QTc prolongation in real life situations; as such calculations have no predictive value, nor can be parameterised for the purposes of non-random effects in clinical trial simulations. Another advantage associated with the use of a PKPD model is that even if the pharmacokinetics has been evaluated at supra-therapeutic levels, drug exposure can be extrapolated to reflect drug concentrations observed after therapeutic doses, so that accurate inferences can be made of the corresponding effect on QT-interval. We also anticipate the importance of our approach to assess the characteristics of high-risk population subgroups by introducing such patients into simulation scenarios.

Despite the potential impact of a tool for "not-in-trial" simulations, the proposed approach in this investigation has some limitations. First, it should be noted that the population estimated parameters rather than the individual estimated parameters were used in the PKPD simulations. This reduces the overall inter-individual variability that exists within a population. In addition, an assumption was made that all the subjects in the Rotterdam cohort were fully compliant with medication consumption so that the same plasma concentration range is reached as those in the clinical trial. QT measurements in the Rotterdam Study cohort were taken during the *d,l*-sotalol prescription. It was therefore assumed that the timing of the QT measurement as well as the differences in measurement equipment contribute to increased random variation without introducing significant bias. Fully automated machine read measurements were used in the Rotterdam cohort, whilst QT-intervals were manually measured by a cardiographist in the clinical trial.

In summary, integration of clinical trial and epidemiological data reveals a flaw in the assumption that findings about drug effects in a clinical trial may be generalisable to the real life population. The assessment of concentration-effect relationships was pivotal for estimating the contribution of drug-induced increase in QTc-intervals. Furthermore, the use of a hierarchical nonlinear mixed effects modelling enabled other causal factors to be incorporated into the evaluation of the overall treatment effect. Our findings show that accurate evaluation of safety requires understanding of treatment response of subjects who are "not-in-trial".

REFERENCES

[1] Shah RR. Drug-induced QT interval prolongation: regulatory perspectives and drug development. Ann Med 2004;36 Suppl 1:47-52.

[2] Yap YG, Camm AJ. Drug induced QT prolongation and torsades de pointes. Heart 2003 Nov;89(11):1363-72.

[3] Dessertenne F. [Ventricular tachycardia with 2 variable opposing foci]. Arch Mal Coeur Vaiss 1966 Feb;59(2):263-72.

[4] Straus SM, Kors JA, De Bruin ML, van der Hooft CS, Hofman A, Heeringa J, et al. Prolonged QTc interval and risk of sudden cardiac death in a population of older adults. J Am Coll Cardiol 2006 Jan 17;47(2):362-7.

[5] Straus SM, Sturkenboom MC, Bleumink GS, Dieleman JP, van der Lei J, de Graeff PA, et al. Non-cardiac QTc-prolonging drugs and the risk of sudden cardiac death. Eur Heart J 2005 Oct;26(19):2007-12.

[6] Darpo B, Nebout T, Sager PT. Clinical evaluation of QT/QTc prolongation and proarrhythmic potential for nonantiarrhythmic drugs: the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use E14 guideline. J Clin Pharmacol 2006 May;46(5):498-507.

[7] Shah RR. Drug-induced prolongation of the QT interval: regulatory dilemmas and implications for approval and labelling of a new chemical entity. Fundam Clin Pharmacol 2002 Apr;16(2):147-56.

[8] Points to Consider: The assessmet of the potential for QT interval prolongation by non-cardiovascular medicinal products. London; 1997.

[9] Note for Guidance on the Investigation of Drug Interactions. London; 1997.

[10] Guidance for Industry: ICH E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrythmic Potential for Non-Antiarrythmic Drugs. U.S. Department of Health and Human Services. Food and Drug Administration (2005)

[11] Chain AS, Krudys KM, Danhof M, Della Pasqua O. Assessing the Probability of Drug-Induced QTc-Interval Prolongation During Clinical Drug Development. Clin Pharmacol Ther 2011 Dec;90(6):867-75.

[12] Garnett CE, Beasley N, Bhattaram VA, Jadhav PR, Madabushi R, Stockbridge N, et al. Concentration-QT relationships play a key role in the evaluation of proarrhythmic risk during regulatory review. J Clin Pharmacol 2008 Jan;48(1):13-8.

[13] Letsas KP, Efremidis M, Kounas SP, Pappas LK, Gavrielatos G, Alexanian IP, et al. Clinical characteristics of patients with drug-induced QT interval prolongation and torsade de pointes: identification of risk factors. Clin Res Cardiol 2009 Apr;98(4):208-12.

[14] Wolbrette DL. Drugs that cause Torsades de pointes and increase the risk of sudden cardiac death. Curr Cardiol Rep 2004 Sep;6(5):379-84.

[15] Bonate PL, Russell T. Assessment of QTc prolongation for non-cardiac-related drugs from a drug development perspective. J Clin Pharmacol 1999 Apr;39(4):349-58.

[16] Innovation or stagnation? Challenge and opportunity on the critical path to new medical products. In: Services DoHaH, editor; 2004.

[17] Kola I, Landis J. Can the pharmaceutical industry reduce attrition rates? Nat Rev Drug Discov 2004 Aug;3(8):711-5.

[18] Chan PL, Holford NH. Drug treatment effects on disease progression. Annu Rev Pharmacol Toxicol 2001;41:625-59.

[19] Gobburu JV, Marroum PJ. Utilisation of pharmacokinetic-pharmacodynamic modelling and simulation in regulatory decision-making. Clin Pharmacokinet 2001;40(12):883-92.

[20] Kimko HHC, Peck CC. Clinical Trial Simulations: Applications and Trends. New York: Springer; 2011.

[21] DiMasi JA, Feldman L, Seckler A, Wilson A. Trends in risks associated with new drug development: success rates for investigational drugs. Clin Pharmacol Ther 2010 Mar;87(3):272-7.

[22] Lalonde RL, Kowalski KG, Hutmacher MM, Ewy W, Nichols DJ, Milligan PA, et al. Model-based drug development. Clin Pharmacol Ther 2007 Jul;82(1):21-32.

[23] Laverty H, Benson C, Cartwright E, Cross M, Garland C, Hammond T, et al. How can we improve our understanding of cardiovascular safety liabilities to develop safer medicines? Br J Pharmacol 2011 Jun;163(4):675-93.

[24] Pater C. Methodological considerations in the design of trials for safety assessment of new drugs and chemical entities. Curr Control Trials Cardiovasc Med 2005 Feb 3;6(1):1.

[25] Pollard CE, Valentin JP, Hammond TG. Strategies to reduce the risk of drug-induced QT interval prolongation: a pharmaceutical company perspective. Br J Pharmacol 2008 Aug;154(7):1538-43.

[26] Choy AM, Darbar D, Dell'Orto S, Roden DM. Exaggerated QT prolongation after cardioversion of atrial fibrillation. J Am Coll Cardiol 1999 Aug;34(2):396-401.

[27] Makkar RR, Fromm BS, Steinman RT, Meissner MD, Lehmann MH. Female gender as a risk factor for torsades de pointes associated with cardiovascular drugs. JAMA 1993 Dec 1;270(21):2590-7.

[28] Nowinski K, Gadler F, Jensen-Urstad M, Bergfeldt L. Transient proarrhythmic state following atrioventricular junction radiofrequency ablation: pathophysiologic mechanisms and recommendations for management. Am J Med 2002 Nov;113(7):596-602.

[29] Torp-Pedersen C, Moller M, Bloch-Thomsen PE, Kober L, Sandoe E, Egstrup K, et al. Dofetilide in patients with congestive heart failure and left ventricular dysfunction. Danish Investigations of Arrhythmia and Mortality on Dofetilide Study Group. N Engl J Med 1999 Sep 16;341(12):857-65.

[30] Black N. Why we need observational studies to evaluate the effectiveness of health care. BMJ 1996 May 11;312(7040):1215-8.

[31] Atkins D. Creating and synthesizing evidence with decision makers in mind: integrating evidence from clinical trials and other study designs. Med Care 2007 Oct;45(10 Supl 2):S16-22.

[32] Hannan EL. Randomized clinical trials and observational studies: guidelines for assessing respective strengths and limitations. JACC Cardiovasc Interv 2008 Jun;1(3):211-7.

[33] Landewe R, van der Heijde D. Primer: challenges in randomized and observational studies. Nat Clin Pract Rheumatol 2007 Nov;3(11):661-6.

[34] Yang W, Zilov A, Soewondo P, Bech OM, Sekkal F, Home PD. Observational studies: going beyond the boundaries of randomized controlled trials. Diabetes Res Clin Pract 2010 May;88 Suppl 1:S3-9.

[35] Hofman A, Grobbee DE, de Jong PT, van den Ouweland FA. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. Eur J Epidemiol 1991 Jul;7(4):403-22.

[36] van Gent CM, van der Voort HA, de Bruyn AM, Klein F. Cholesterol determinations. A comparative study of methods with special reference to enzymatic procedures. Clin Chim Acta 1977 Mar 1;75(2):243-51.

[37] 1999 World Health Organization-International Society of Hypertension Guidelines for the Management of Hypertension. Guidelines Subcommittee. J Hypertens 1999 Feb;17(2):151-83.

[38] Diabetes Mellitus. Technical Reports series 894. Geneva; 1985.

[39] Bots ML, Hoes AW, Koudstaal PJ, Hofman A, Grobbee DE. Common carotid intima-media thickness and risk of stroke and myocardial infarction: the Rotterdam Study. Circulation 1997 Sep 2;96(5):1432-7.

[40] Bleumink GS, Knetsch AM, Sturkenboom MC, Straus SM, Hofman A, Deckers JW, et al. Quantifying the heart failure epidemic: prevalence, incidence rate, lifetime

risk and prognosis of heart failure The Rotterdam Study. Eur Heart J 2004 Sep;25(18):1614-9.

[41] Mosterd A, Hoes AW, de Bruyne MC, Deckers JW, Linker DT, Hofman A, et al. Prevalence of heart failure and left ventricular dysfunction in the general population; The Rotterdam Study. Eur Heart J 1999 Mar;20(6):447-55.

[42] International classification of diseases and health related problems. Geneva; 1992.

[43] Beal S, Sheiner, L.NONMEM user's guide NONMEM project group. San Francisco; 1999.

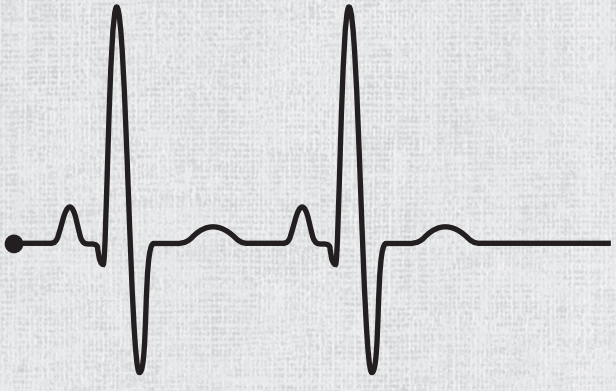
[44] Lunn DJ, Thomas, A., Best, N., Spiegelhalter, D. WinBUGS - A Bayesian modelling framework: concepts, structure, and extensibility. Stat Comput 2000;10:325-37.

[45] de Bruyne MC, Kors JA, Hoes AW, Kruijssen DA, Deckers JW, Grosfeld M, et al. Diagnostic interpretation of electrocardiograms in population-based research: computer program research physicians, or cardiologists? J Clin Epidemiol 1997 Aug;50(8):947-52.

[46] Willems JL, Abreu-Lima C, Arnaud P, van Bemmel JH, Brohet C, Degani R, et al. The diagnostic performance of computer programs for the interpretation of electrocardiograms. N Engl J Med 1991 Dec 19;325(25):1767-73.

[47] Bazett HC. The time relations of the blood-pressure changes after excision of the adrenal glands, with some observations on blood volume changes. J Physiol 1920 Feb 20;53(5):320-39.

[48] Cox DR. Ression Models and Life-Tables. Journal of the Royal Statistical Society Series B (Methodological) 1972;34:187-220.

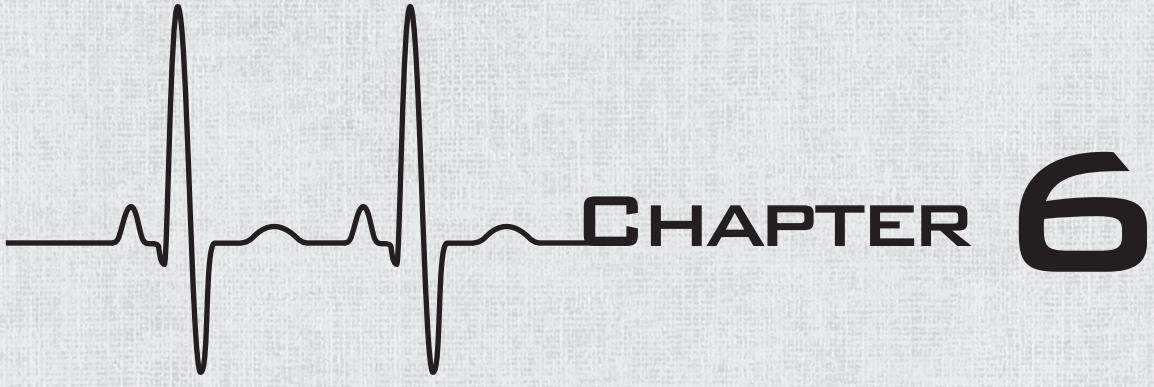


Not-In-Trial Simulation II:
Predicting Cardiovascular Risk From Clinical Trial Data

ASY Chain (1, 2), JP Dieleman (2), M Danhof (1),
MCJM Sturkenboom (2, 3), OE Della Pasqua (1,4)

(1) Leiden/Amsterdam Centre for Drug Research,
Division of Pharmacology, Leiden University, Leiden, The Netherlands
(2) Department of Medical Informatics and Epidemiology,
Erasmus Medical Centre, The Netherlands
(3) Department of Epidemiology, Erasmus Medical Centre,
The Netherlands
(4) Clinical Pharmacology and Discovery Medicine,
GlaxoSmithKline, Stockley Park, Uxbridge, United Kingdom

Submitted for publication



Chapter 6

ABSTRACT

Model-based techniques have become an established approach in drug development. Less attention however has been paid to the role of simulations for the evaluation of safety data, which is fraught with ethical and clinical hurdles. Here we materialise the implementation of Not-In-Trial Simulations as an inferential tool for predicting and translating drug-induced effects from clinical settings to real life situations. We illustrate the concept using QTc-interval prolongation as example of safety concern for both industry and regulatory agencies. In a previous work, we have demonstrated that observed QTc prolongation in a patient population can be only partially explained by the underlying treatment effects. Using a series of simulations, the current investigation shows how QTc-intervals can be modelled as a function of baseline QT-interval, drug exposure, concomitant medications and comorbidity conditions. Our method offers the basis for improved risk management, dissecting drug-specific contribution to pre- or post-approval safety findings.

6.1 INTRODUCTION

Generating safety data that reflects the actual risk (and consequently risk:benefit ratio) for the target population during the course of drug development is fraught with practical and ethical hurdles, particularly if safety concerns regard low frequency or rare events. Such information can only be derived from epidemiological or pharmacoepidemiological studies, which are planned and performed after drug approval. Furthermore, inclusion and exclusion criteria are imposed on clinical trial protocols to mitigate risk and prevent the most vulnerable patients from exposure to an experimental agent for which the risk:benefit ratio is unknown at the time of the investigation. Subsequently, however, drug prescription is not restricted or contra-indicated for those patients who were excluded during the clinical development phase. Implicitly, such practice imposes the assumption that such inclusion/exclusion criteria do not alter treatment outcome. Inferential methods offer an opportunity to address this issue in a more quantitative and systematic manner.

With the exception of signal detection methodologies used for pharmacovigilance purposes [1, 2], the integrated use of clinical trial and epidemiological data for the prediction and interpretation of safety findings in clinical drug development remains very limited [3, 4]. Simulation techniques can play an important role in the implementation of such methods. Unfortunately, the use of simulations has been rather limited in the evaluation of the safety profile whether in epidemiological cohorts or in randomised clinical trials [5].

Using a model-based approach that allows the integration of pharmacokinetic-pharmacodynamic relationships for the evaluation of drug-induced QTc-interval prolongation, we illustrate the concept of Not-In-Trial Simulation (NITS) as an inferential tool for the quantification of the contribution of intrinsic and extrinsic factors on the risk of pro-arrhythmic events, i.e. QTc-interval prolongation. In contrast to typical clinical trial simulations [6-8], we make use of simulations to characterise the role of design factors which have been omitted or excluded from a randomised trial, among which those listed under the exclusion criteria or diverging from the clinical conditions relevant for the target population. Thus, our approach will represent a natural extension of ongoing efforts within the pharmaceutical industry to improve safety signal detection [9-15]. This

approach provides the pharmacological basis for the assessment of causality, discriminating drug-induced from other (drug-unrelated) effects.

The objective of this study was therefore to show how clinical findings translate into real life observations and resolve some of the discrepancies commonly observed between pre- vs. post-market estimates of the pro-arrhythmic effects associated with a new therapeutic agent. Drug-induced QTc-interval prolongation has become one of the most widely debated topics in the past decade [16-19], given its presumed association with increased risk for ventricular arrhythmias, which can potentially lead to sudden cardiac death [20, 21]. *d,l*-Sotalol was selected as a paradigm compound for the purposes of our investigation. Based on simulation scenarios we explain the impact and clinical implications of additional factors on QTc-interval prolongation for the real life population.

Despite numerous efforts aimed at improving signal detection of cardiovascular events for new medicines [22-24], none of them has focused on what actually happens after the drug has been approved and released into the general population. Although many post-marketing surveillance trials and spontaneous adverse events reporting have been used to monitor the incidence of (supra)ventricular arrhythmias, torsade de pointes and other safety events [25-27], there is still an important advantage to knowing what can be expected so that mitigation plans can be made in advance. Thus, we anticipate that our approach will widen the views on risk management beyond the evolving perspective from clinical pharmacology experts and regulators, i.e., that the liability for QTc-interval prolongation cannot be assessed accurately without an assessment of the concentration-effect relationships [28]. Not-In-Trial Simulations will enable quantitative evaluation of the implication of other factors contributing to QTc-interval prolongation in real life population, as compared to the observed drug effects in clinical trials. Among various factors [15], our example reveals how gender [29], congestive heart failure [30] and baseline QT-intervals [31] can have equally important impact on treatment outcome.

6.2 METHODS

Evaluation of the effect of age on QTc-interval

Baseline information from seven healthy volunteer studies was used for the development of the relationship between baseline QTc-intervals and age. These data were combined with elderly patient data that were classified as healthy in the Rotterdam Study cohort described below. The healthy elderly subjects were defined as those who were not on any medications and not have any comorbidity conditions recorded at the time of the ECG measurement. Only the Bazett correction factor was used for the assessment of the effect of age on QT-interval due to the consistency availability of data across all datasets. Other correction factors were not explored nor did we explore the effect of age on RR-intervals. Finally, a quadratic regression was used to describe the age-dependent QTc data. Based on the resulting parameter estimates, the regression was reduced to a linear relationship in R2.8 (The R project - http://www.r-project.org/).

Evaluation of drug-induced effects in healthy volunteers

Data from a Phase I healthy volunteer study with *d,l*-sotalol was selected from GlaxoSmithKline's clinical data repository (study number EXP20001). The study was performed according to a placebo-controlled, three-way crossover design with two placebo periods. The study population comprised of 30 subjects (12 females, 18 males), between the ages of 19 and 47. Other relevant eligibility criteria for inclusion into the study for the purposes of our investigation were non-smoking status, body mass index between 19-30 kg/m^2, with a weight of 50 – 95 kg, inclusive. Subjects with a history of hypertension, asthma, bronchial hyperactivity or peripheral vascular disease, including diabetes or previous alcohol or drug abuse were excluded at screening. In addition, subjects showing abnormal QRST complex morphology, sinus bradycardia (HR<45 bmp, PR > 210 msec) or QTc-interval values above 420 or 440 msec (for male and female, respectively) were also excluded.

Modelling of QTc-interval prolongation:

The primary outcome of this study was to demonstrate the use of the Not-In-Trial Simulation concept as a tool to quantify the contribution of additional

factors to QTc-interval prolongation. The simulation of QTc-intervals was performed in a two-step approach. First, drug exposure in the target population (i.e., *d,l*-sotalol users) were derived using a pharmacokinetic model previously published [37]. Briefly, the pharmacokinetics of *d, l*-sotalol was described by a two compartment, first-order absorption model with weight as a covariate on clearance using nonlinear mixed effects modelling, as implemented in NONMEM (v5.1, ICON Development Solutions**,** Ellicott City, MD, USA) [38].

Heart-rate corrected QT-intervals were described by a pharmacokinetic-pharmacodynamic (PKPD) model using a Bayesian approach in which physiological (QT and RR) and drug specific (concentration) parameters are parameterised in an independent manner (Equation 1) [39]. A detailed description of model development was presented in a previous publication [37].

$$QT_c = QT_0 \cdot RR^{\alpha} + A \cdot \cos\left(\frac{2\pi}{24}(t - \phi)\right) + slope \cdot C$$

Equation 1

where, QT_0 [msec] is the intercept of the QT-RR relationship. Sex was included as a covariate for this parameter whenever applicable. RR [sec] is the interval between successive R waves, α is the individual heart rate correction factor, A [msec] is the amplitude of circadian rhythm, t is the clock time, φ is the phase, slope [msec/concentration] is the linear pharmacodynamic relationship, and C is the predicted concentration of drug at the time of QT-interval measurements.

Drug-induced QTc-intervals in patients was simulated assuming full compliance to treatment. Population parameter estimates were used for the simulations (slope = 0.02 msec/ng/ml, and intercept (QT_0) = 380 msec).

Evaluation of effect of intrinsic and extrinsic factors on QTc-interval in real life patients

Data from a real life population setting was obtained from the Rotterdam Study. The study is a prospective population-based cohort study, which started with a baseline visit between 1990 and 1993. All inhabitants of a suburb in Rotterdam, Ommoord, aged 55 years and over were invited to participate. Of the 10 275 subjects invited, 7983 (78%) gave their written informed consent and took part in the baseline examination. The study was approved by the Medical Ethics Committee of the Erasmus Medical Center, Rotterdam, The Netherlands. Objectives and methods of investigation have been described in detail in other publications [40]. All participants were visited at home for a standardised questionnaire and 7151 were subsequently examined at the research center. A second, third and fourth follow up visit took place, respectively, between 1993 - 1995, 1997 - 1999 and 2002 - 2004. In addition to follow-up examinations, the cohort is continuously monitored for major morbidity and mortality through linkage with general practitioner and municipality records.

All subjects who started treatment with *d,l*-sotalol during follow-up were included in the present study, with the exception of prevalent users and subjects with left ventricular hypertrophy (LVH), right bundle-branch block (RBBB) and left bundle-branch block (LBBB), who were excluded. The reason for the exclusion of these patients is that LVH, RBBB and LBBB can cause secondary repolarisation changes and atrial fibrillation can cause difficulties in measuring QT-intervals. From the remaining population, "new" *d,l*-sotalol users were identified as those receiving their first prescription after enrolment into the study. Our analysis included all QTc measurements after the start of *d,l*-sotalol for all new users.

Comorbidities were screened and identified according to standard clinical criteria. Hypertension was defined as systolic blood pressure > 160 mm Hg or diastolic blood pressure > 100 mm Hg and/or use of antihypertensive medications, encompassing grade 2 and grade 3 hypertension compounds according to the World Health Organisation (WHO) criteria [41]. Diabetes mellitus was defined as the use of blood glucose-lowering medication and / or a non-fasting or post-load serum glucose level of 11.1 nmol/L or higher according to the WHO [42]. A history of myocardial infarction was assessed by self-report, checked with records from the general practitioner or

cardiologist and/or electrocardiographic evidence. All reported myocardial infarctions were verified against the medical records as described in detail previously [43]. Assessment of heart failure at baseline and during follow-up was assessed by reviewing all medical records for the occurrence of at least two signs and symptoms suggestive of heart failure or the use of medication for the indication heart failure and hospital discharge letters. Cases of incident heart failure were obtained by continuously monitoring the participants [44, 45]. The ankle arm index (AAI) was used as potential predictor of cardiovascular diseases and mortality [46].

The impact of co-medications and comorbidities were evaluated using a linear regression in R2.8 where the absolute differences in QTc prolongation between taking *d,l*-sotalol alone and in conjunction with co-medications and comorbidities were calculated with standard deviation. The same evaluation was performed for non-users.

ECG measurement in clinical trials and real life population

A 10 sec 12-lead ECG, resulting on average 8-10 heart beats, was recorded with an ACTA electrocardiograph (ESAOTE, Florence, Italy) at a sampling frequency of 500 Hz and stored digitally. All ECGs were processed by the Modular ECG Analysis System (MEANS) to obtain ECG measurements. The MEANS program has been evaluated extensively and reported elsewhere [47-49]. MEANS determines common onsets and offsets for all 12 leads together for one representative averaged beat, with the use of template matching techniques [48], until the end of the T wave. To adjust for heart rate, Bazett's formula (QTc = QT / RR ^ ½) was used [50].

Not-In-Trial Simulation

The final distribution of QTc values associated with all causal factors was simulated stepwise in R2.8. Differences between observed and expected QTc-intervals were compared non-parametrically by visual inspection.

6.3 RESULTS

From the Rotterdam Study cohort, 158 "new" *d,l*-sotalol users were identified which was defined as only those who have started a prescription

for the drug after the beginning of the study in 1990. The final simulated QTc-interval (expected) distribution of the real life *d,l*-sotalol users is composed of four parts. It is a cumulative function of age-dependent baseline QTc values, drug effects from *d,l*-sotalol, contribution from concomitant medication and comorbidity conditions as well as a between-subject variability term. The Bazett formula has been used throughout this study for the heart rate-corrected QT values as this was a consistent output between various data used for the investigation.

Age-dependent baseline QTc values

Using all the available baseline QTc values for healthy subjects in this study, a linear relationship was found to best describe the data. The results are also separated by gender because of the known differences in cardiac outputs between men and women [32]. The correlation between age and baseline QTc was modelled according to equations: QTc = 0.5096(Age) + 386.52 for the male population and QTc = 0.8617(Age) + 371.34 for the female population as captured in Figure 1.

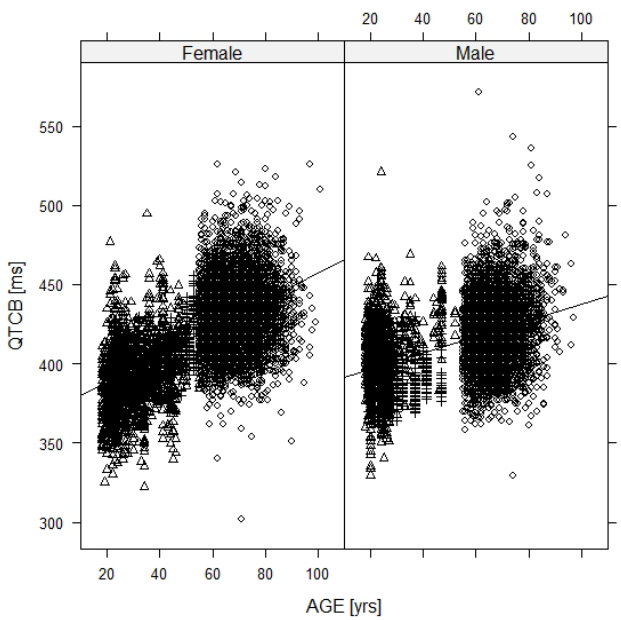


Figure 1. The relationship between age and healthy baseline QTc values, separated by gender.

PKPD modelling of QTc-interval prolongation

Pharmacokinetics (PK) was modelled separately where it was found that a two compartment model best described the time course of concentrations in plasma. The absorption phase was characterised by a lag time, and furthermore, it was found that body weight had a significant effect on the volume of distribution. The PKPD model, shown in Figure 2, allowed the system-specific parameters (i.e., those related to the individual correction for the QT-RR relationship and circadian rhythm) to be estimated separately from the drug-specific parameters (i.e., drug-induced QTc-interval prolongation).

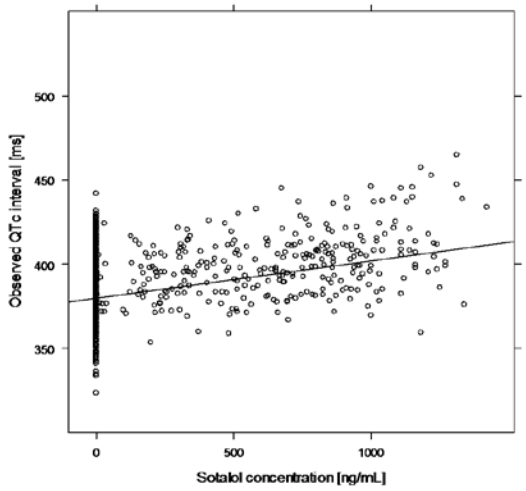


Figure 2. Pharmacokinetic-pharmacodynamic relationship of d,l-sotalol in healthy subjects following administration of 160 mg.

Effects of concomitant medications and comorbidity conditions

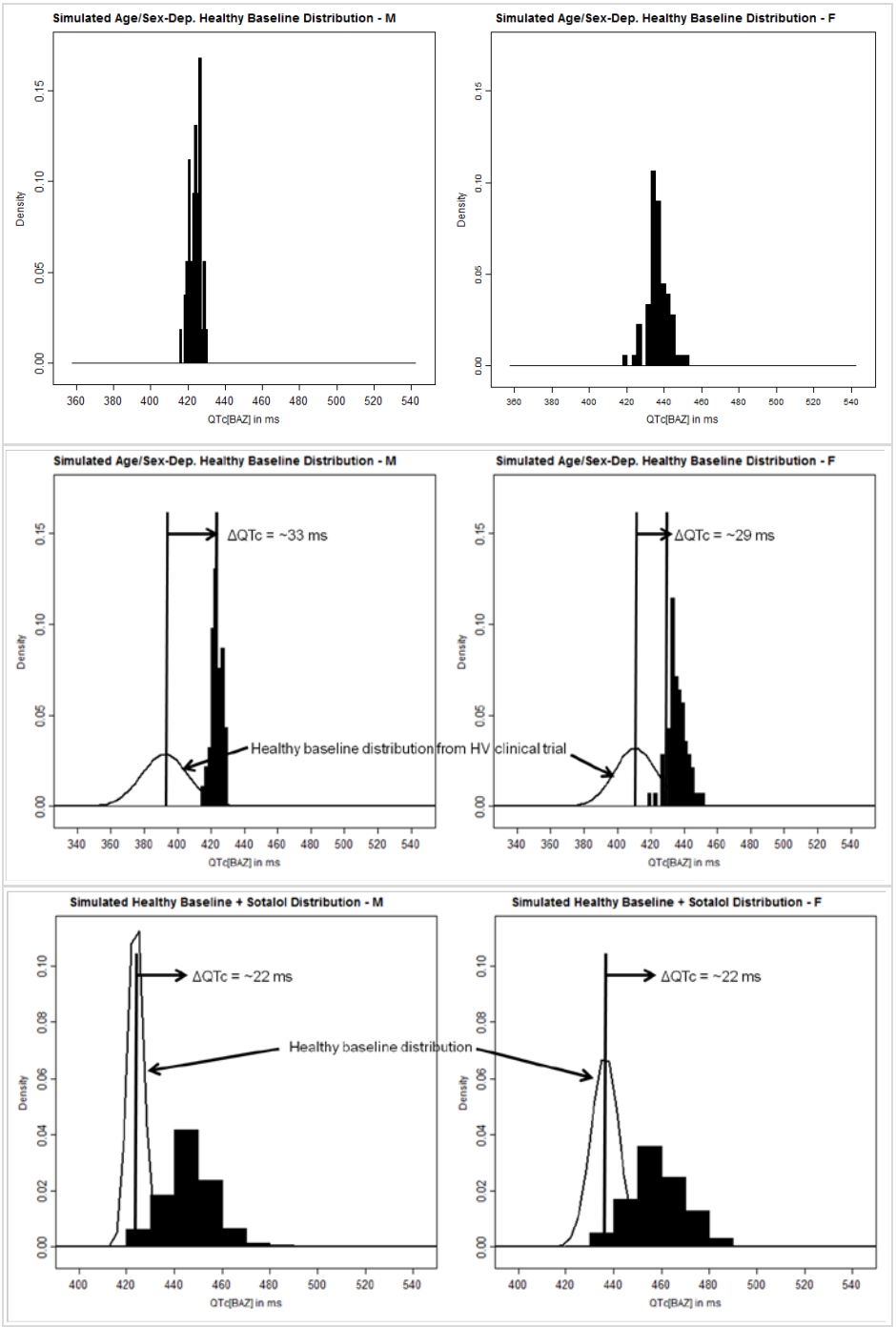
Linear regression calculations assessing the impact of various concomitant medications and comorbidity conditions revealed different increase of QTc values in milliseconds. There was also no interactions found between *d,l*-sotalol and any of the conditions, thus the results found with the users and non-users were not statistically different. Since the power for the non-user group was greater, the results were used in the simulation and the increase in QTc values are summarised in Table 1.

	Increase in QTc [msec]	
	Male, (CI)	Female, (CI)
Heart Failure	7.7 (+/- 4.02)	5.8 (+/- 3.49)
Myocardial Infarction	3 (+/- 2.11)	3.2 (+/- 2.31)
Arrhythmia	18.5 (+/- 9.24)	20.4 (+/- 9.51)
Diabetes	3.0 (+/- 2.28)	5.4 (+/- 2.0)

Table 1. Increase in QTc [msec] due to various concomitant medications and comorbidity conditions

Not-In-Trial Simulation

The newly proposed Not-In-Trial Simulation model yielded the results shown in Figures 3a-f. First using the ages of the 158 real life *d,l*-sotalol users in the Rotterdam Study, the baseline QTc values were simulated (Figure 3a) using age-dependent relationships shown above. In Figure 3b, the baseline QTc distributions of the real life population is compared with the baseline distributions from the Phase I clinical trial of healthy volunteers. From the comparison, it was found that there were discrepancies of 33 msec and 29 msec in male and female respectively. Then the effect of *d,l*-sotalol was added to the baseline distributions of the real life population. The QTc values were shifted by 22 msec in both genders in Figure 3c. Afterwards, taking the comorbidity and concomitant medications of the real life *d,l*-sotalol users were taken into consideration, the mean QTc values were increased by 3 msec in the male population and 1 msec in the female population (Figure 3d). In Figure 3e, the between-subject variability boundaries around the QTc distributions were added, which has until now been constructed using only the mean values from each step. Finally in Figure 3f, the observed QTc values from the 158 *d,l*-sotalol users were superimposed onto the simulated distributions. It can be visually confirmed that real life observed QTc measurements fell within the boundaries of the final simulated QTc distributions along with the between-subject variability found in the age-dependent baseline values.


Simulated Age/Sex-Dep. Healthy Baseline Distribution - M
Simulated Age/Sex-Dep. Healthy Baseline Distribution - F
Density
QTc[BAZ] in ms
ΔQTc = ~33 ms
ΔQTc = ~29 ms
Healthy baseline distribution from HV clinical trial
Simulated Healthy Baseline + Sotalol Distribution - M
Simulated Healthy Baseline + Sotalol Distribution - F
ΔQTc = ~22 ms
Healthy baseline distribution

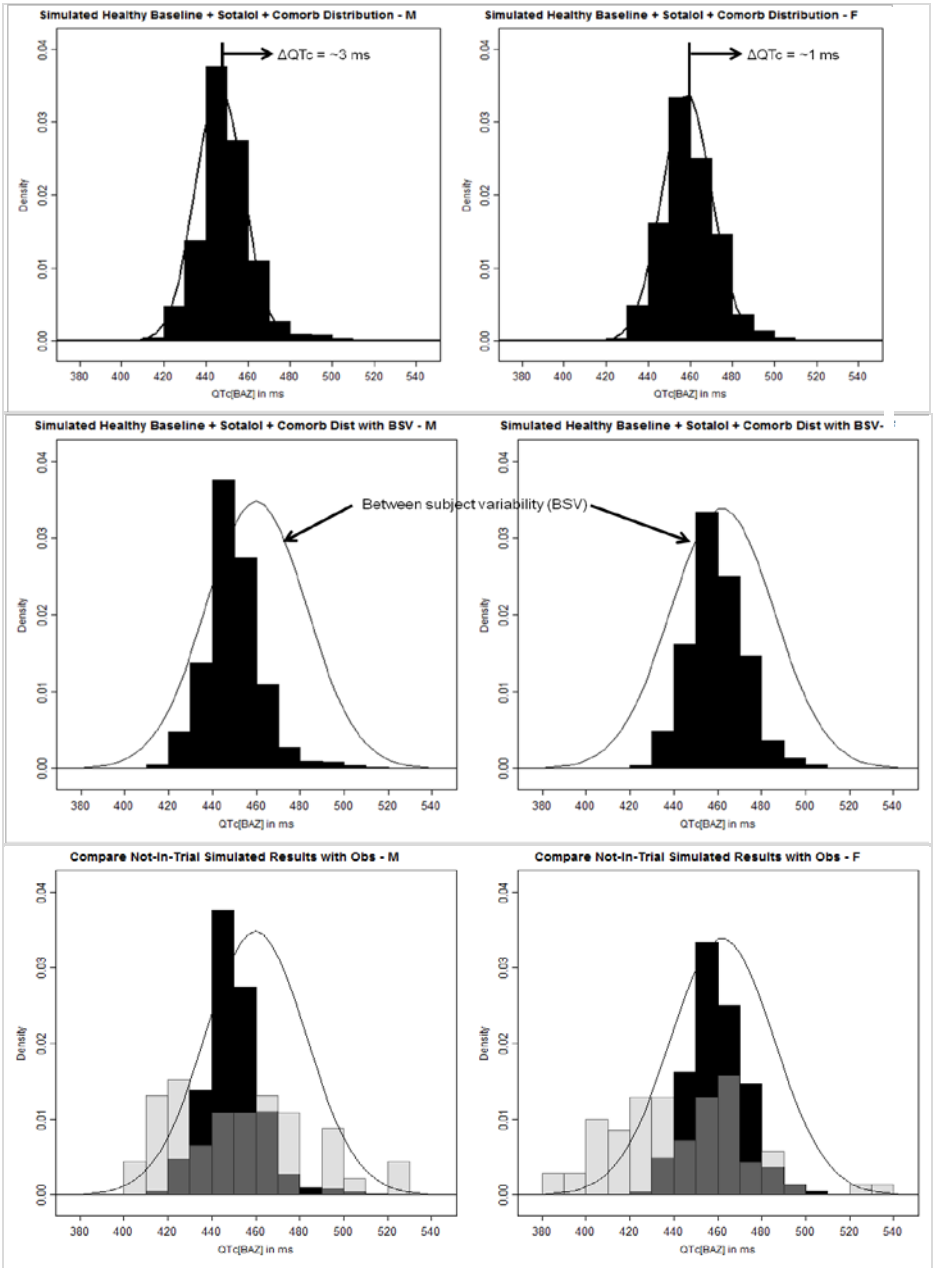


Figure 3. Overall simulation model results (a-f). The left panel depicts results for the male population and female outcome is shown on the right. The lighter colours represent the actual observed QTc values as recorded in the Rotterdam Study while the simulated results are superimposed in a darker colour. The Gaussian distribution curve represents the lower and upper bound of the between-subject variability in baseline values.

6.4 DISCUSSIONS

In contrast to current practice which relies on the evidence from a thorough QTc study for the assessment of the risk of pro-arrhythmic events, we present a model-based approach for the characterisation of drug-induced effects which enables explicit assessment of causality, distinguishing drug-related effects from other factors contributing to the QTc-interval prolongation, namely that, QTc (real life population) = age-dependent baseline QTc values + fluctuation in QTc due to drug effects + effects on QTc due to concomitant medications and comorbidity conditions.

Although QT-interval prolongation is not necessarily harmful, it is at present perhaps the best and only surrogate marker for torsade de pointes, especially so because the definition of this arrhythmia requires prolongation of QT-interval as a preceding event. The crucial question from regulatory perspective is "How efficient and reliable are the pre-approval clinical trials in identifying the clinical risk of torsade de pointes, given the 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" [33]. For a variety of reasons, clinical trials may or may not detect the frequency and intensity of QT-interval prolongation but they are most unlikely to detect the risk of torsade de pointes. Firstly, safety trial designs are highly efficacy-oriented. Secondly, the number of patients exposed to the drug in the pre-approval phase is not large enough to detect a relatively rare but potentially fatal risk. Thirdly, many subgroups of patients most at risk of torsade de pointes during the uncontrolled clinical use to the drug in question are usually excluded from these trials. These subgroups include: females; the elderly, those with predisposing cardiac or noncardiac 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 the potassium channels; those susceptible to bradycardia or electrolyte imbalance; or those receiving drugs with a potential for pharmacokinetic or pharmacodynamic interactions. Therefore the scope for detecting drug-drug or drug-disease interactions in clinical trials is also very limited. And yet, experience has shown that these are among the most important risk factors.

Based on the aforementioned, one may question whether conducting the thorough-QT (TQT) studies yield the best indicator of risk for new therapeutic agents. This regulatory practice is further challenged by increasing clinical evidence that there is no fixed relationship between the extent of QT-interval prolongation and the risk of a patient developing torsade de pointes [34].

Our findings show that baseline QTc values are age and gender dependent. This correlation plays an important role in that observed QTc values will increase over time irrespective of drug treatment. Drug effect and comorbidity conditions also contribute to further increase in QTc values. In this study, simulation scenarios were implemented in such a way that all relevant causal factors affecting QTc prolongation are combined in an integrated manner. It is envisioned that this approach may provide more accurate estimates of cardiovascular risks and enable mitigation of these risks in the target population, preventing future drug-withdrawal due to cardiovascular safety.

Limitations

In this study we have successfully demonstrated the concept of Not-In-Trial Simulation. However, there are limitations in this present study that should be mentioned as this is the first attempt at realising the idea of using simulation as a technique to translate real life observations using clinical trial findings. First and foremost is the lack of hierarchical modeling techniques used in estimating the relationship between age and QTc values as well as the increase due to comorbidity conditions and concomitant drugs which can affect the accuracy or precision of effect sizes and the parameter estimations. Currently, we can verify that our simulation encompasses the range of values in real life situations, but to be able to improve the precision of the parameter estimations would lead to more confidence when performing the simulations for future compounds prospectively.

Another limitation to this study is that the age-QTc relationships in healthy subjects as well as the contributions from comorbidities and co-medications are limited to only the datasets used presently. This approach would no doubt benefit from a larger database to improve all the parameter estimates.

Model-based risk management

Currently, both the US and Europe regulatory agencies have produced guidelines that outline efforts in managing the safety and risk of pre- and post-market medical products [35, 36]. Enhancing internal and external epidemiologic and informatics capabilities, strengthening methods and tools of safety surveillance, developing and incorporating new quantitative tools in the assessment of benefit and risk are examples of ongoing activities.

To date, recommendations on how best to detect safety signals such as the introduction of the thorough-QT studies are being published. However, utilising modeling and simulation techniques as a quantitative tool to assess clinical relevance prior to implementation of mitigation measures or eventually as basis for appropriate recommendations is still not adequately advocated.

REFERENCES

[1] Hauben M, Aronson JK. Defining 'signal' and its subtypes in pharmacovigilance based on a systematic review of previous definitions. Drug Saf 2009;32(2):99-110.

[2] Hauben M, Zhou X. Quantitative methods in pharmacovigilance: focus on signal detection. Drug Saf 2003;26(3):159-86.

[3] Atkins D. Creating and synthesizing evidence with decision makers in mind: integrating evidence from clinical trials and other study designs. Med Care 2007 Oct;45(10 Supl 2):S16-22.

[4] Hannan EL. Randomized clinical trials and observational studies: guidelines for assessing respective strengths and limitations. JACC Cardiovasc Interv 2008 Jun;1(3):211-7.

[5] Narukawa M, Yafune A, Takeuchi M. Observation of time-dependent adverse events and the influence of drop-out thereon in long-term safety studies--simulation study under the current practice of post-marketing safety evaluation in Japan. J Biopharm Stat 2004 May;14(2):403-14.

[6] Chan PL, Holford NH. Drug treatment effects on disease progression. Annu Rev Pharmacol Toxicol 2001;41:625-59.

[7] Gobburu JV, Marroum PJ. Utilisation of pharmacokinetic-pharmacodynamic modelling and simulation in regulatory decision-making. Clin Pharmacokinet 2001;40(12):883-92.

[8] Holford NH, Kimko HC, Monteleone JP, Peck CC. Simulation of clinical trials. Annu Rev Pharmacol Toxicol 2000;40:209-34.

[9] DiMasi JA, Feldman L, Seckler A, Wilson A. Trends in risks associated with new drug development: success rates for investigational drugs. Clin Pharmacol Ther Mar;87(3):272-7.

[10] Fermini B, Fossa AA. The impact of drug-induced QT interval prolongation on drug discovery and development. Nat Rev Drug Discov 2003 Jun;2(6):439-47.

[11] Lalonde RL, Kowalski KG, Hutmacher MM, Ewy W, Nichols DJ, Milligan PA, et al. Model-based drug development. Clin Pharmacol Ther 2007 Jul;82(1):21-32.

[12] Laverty H, Benson C, Cartwright E, Cross M, Garland C, Hammond T, et al. How can we improve our understanding of cardiovascular safety liabilities to develop safer medicines? Br J Pharmacol Jun;163(4):675-93.

[13] Pater C. Methodological considerations in the design of trials for safety assessment of new drugs and chemical entities. Curr Control Trials Cardiovasc Med 2005 Feb 3;6(1):1.

[14] Pollard CE, Valentin JP, Hammond TG. Strategies to reduce the risk of drug-induced QT interval prolongation: a pharmaceutical company perspective. Br J Pharmacol 2008 Aug;154(7):1538-43.

[15] Roden DM. Drug-induced prolongation of the QT interval. N Engl J Med 2004 Mar 4;350(10):1013-22.

[16] Darpo B, Nebout T, Sager PT. Clinical evaluation of QT/QTc prolongation and proarrhythmic potential for nonantiarrhythmic drugs: the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use E14 guideline. J Clin Pharmacol 2006 May;46(5):498-507.

[17] Ahmad K, Dorian P. Drug-induced QT prolongation and proarrhythmia: an inevitable link? Europace 2007 Sep;9 Suppl 4:iv16-22.

[18] De Ponti F, Poluzzi E, Montanaro N. QT-interval prolongation by non-cardiac drugs: lessons to be learned from recent experience. Eur J Clin Pharmacol 2000 Apr;56(1):1-18.

[19] Shah RR. Drug-induced QT interval prolongation: regulatory perspectives and drug development. Ann Med 2004;36 Suppl 1:47-52.

[20] Dessertenne F. [Ventricular tachycardia with 2 variable opposing foci]. Arch Mal Coeur Vaiss 1966 Feb;59(2):263-72.

[21] Yap YG, Camm AJ. Drug induced QT prolongation and torsades de pointes. Heart 2003 Nov;89(11):1363-72.

[22] Haverkamp W, Breithardt G, Camm AJ, Janse MJ, Rosen MR, Antzelevitch C, et al. The potential for QT prolongation and proarrhythmia by non-antiarrhythmic drugs: clinical and regulatory implications. Report on a policy conference of the European Society of Cardiology. Eur Heart J 2000 Aug;21(15):1216-31.

[23] Netzer R, Ebneth A, Bischoff U, Pongs O. Screening lead compounds for QT interval prolongation. Drug Discov Today 2001 Jan 1;6(2):78-84.

[24] Shah RR, Hondeghem LM. Refining detection of drug-induced proarrhythmia: QT interval and TRIaD. Heart Rhythm 2005 Jul;2(7):758-72.

[25] de Bruyne MC, Hoes AW, Kors JA, Hofman A, van Bemmel JH, Grobbee DE. Prolonged QT interval predicts cardiac and all-cause mortality in the elderly. The Rotterdam Study. Eur Heart J 1999 Feb;20(4):278-84.

[26] Dekker JM, Schouten EG, Klootwijk P, Pool J, Kromhout D. Association between QT interval and coronary heart disease in middle-aged and elderly men. The Zutphen Study. Circulation 1994 Aug;90(2):779-85.

[27] Montanez A, Ruskin JN, Hebert PR, Lamas GA, Hennekens CH. Prolonged QTc interval and risks of total and cardiovascular mortality and sudden death in the general population: a review and qualitative overview of the prospective cohort studies. Arch Intern Med 2004 May 10;164(9):943-8.

[28] Garnett CE, Beasley N, Bhattaram VA, Jadhav PR, Madabushi R, Stockbridge N, et al. Concentration-QT relationships play a key role in the evaluation of proarrhythmic risk during regulatory review. J Clin Pharmacol 2008 Jan;48(1):13-8.

[29] Makkar RR, Fromm BS, Steinman RT, Meissner MD, Lehmann MH. Female gender as a risk factor for torsades de pointes associated with cardiovascular drugs. JAMA 1993 Dec 1;270(21):2590-7.

[30] Torp-Pedersen C, Moller M, Bloch-Thomsen PE, Kober L, Sandoe E, Egstrup K, et al. Dofetilide in patients with congestive heart failure and left ventricular dysfunction. Danish Investigations of Arrhythmia and Mortality on Dofetilide Study Group. N Engl J Med 1999 Sep 16;341(12):857-65.

[31] Houltz B, Darpo B, Edvardsson N, Blomstrom P, Brachmann J, Crijns HJ, et al. Electrocardiographic and clinical predictors of torsades de pointes induced by almokalant infusion in patients with chronic atrial fibrillation or flutter: a prospective study. Pacing Clin Electrophysiol 1998 May;21(5):1044-57.

[32] Kassotis J, Costeas C, Bedi AK, Tolat A, Reiffel J. Effects of aging and gender on QT dispersion in an overtly healthy population. Pacing Clin Electrophysiol 2000 Jul;23(7):1121-6.

[33] Bonate PL, Russell T. Assessment of QTc prolongation for non-cardiac-related drugs from a drug development perspective. J Clin Pharmacol 1999 Apr;39(4):349-58.

[34] Shah RR. Pharmacogenetic aspects of drug-induced torsade de pointes: potential tool for improving clinical drug development and prescribing. Drug Saf 2004;27(3):145-72.

[35] The Future of Drug Safety — Promoting and Protecting the Health of the Public. In: Administration FaD, editor; 2007.

[36] Guideline on risk management systems for medicinal products for human use. 2005.

[37] Chain AS, Krudys KM, Danhof M, Della Pasqua O. Assessing the Probability of Drug-Induced QTc-Interval Prolongation During Clinical Drug Development. Clin Pharmacol Ther Dec;90(6):867-75.

[38] Beal S, Sheiner, L. . NONMEM user's guide NONMEM project group. San Francisco; 1999.

[39] Lunn DJ, Thomas, A., Best, N., Spiegelhalter, D. WinBUGS - A Bayesian modelling framework: concepts, structure, and extensibility. Stat Comput 2000;10:325-37.

[40] Hofman A, Grobbee DE, de Jong PT, van den Ouweland FA. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. Eur J Epidemiol 1991 Jul;7(4):403-22.

[41] 1999 World Health Organization-International Society of Hypertension Guidelines for the Management of Hypertension. Guidelines Subcommittee. J Hypertens 1999 Feb;17(2):151-83.

[42] Diabetes Mellitus. Technical Reports series 894. Geneva; 1985.

[43] Bots ML, Hoes AW, Koudstaal PJ, Hofman A, Grobbee DE. Common carotid intima-media thickness and risk of stroke and myocardial infarction: the Rotterdam Study. Circulation 1997 Sep 2;96(5):1432-7.

[44] Bleumink GS, Knetsch AM, Sturkenboom MC, Straus SM, Hofman A, Deckers JW, et al. Quantifying the heart failure epidemic: prevalence, incidence rate, lifetime risk and prognosis of heart failure The Rotterdam Study. Eur Heart J 2004 Sep;25(18):1614-9.

[45] Mosterd A, Hoes AW, de Bruyne MC, Deckers JW, Linker DT, Hofman A, et al. Prevalence of heart failure and left ventricular dysfunction in the general population; The Rotterdam Study. Eur Heart J 1999 Mar;20(6):447-55.

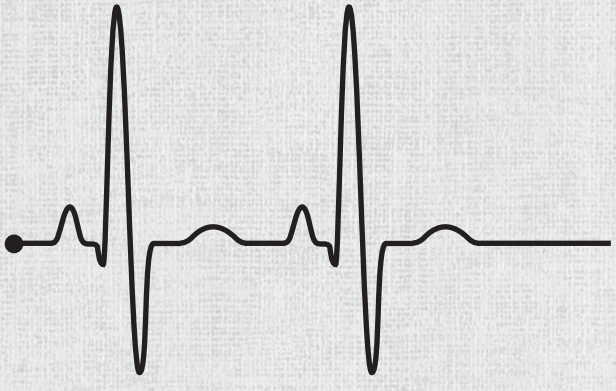
[46] International classification of diseases and health related problems. Geneva; 1992.

[47] de Bruyne MC, Kors JA, Hoes AW, Kruijssen DA, Deckers JW, Grosfeld M, et al. Diagnostic interpretation of electrocardiograms in population-based research: computer program research physicians, or cardiologists? J Clin Epidemiol 1997 Aug;50(8):947-52.

[48] van Bemmel JH, Kors JA, van Herpen G. Methodology of the modular ECG analysis system MEANS. Methods Inf Med 1990 Sep;29(4):346-53.

[49] Willems JL, Abreu-Lima C, Arnaud P, van Bemmel JH, Brohet C, Degani R, et al. The diagnostic performance of computer programs for the interpretation of electrocardiograms. N Engl J Med 1991 Dec 19;325(25):1767-73.

[50] Bazett HC. The time relations of the blood-pressure changes after excision of the adrenal glands, with some observations on blood volume changes. J Physiol 1920 Feb 20;53(5):320-39.



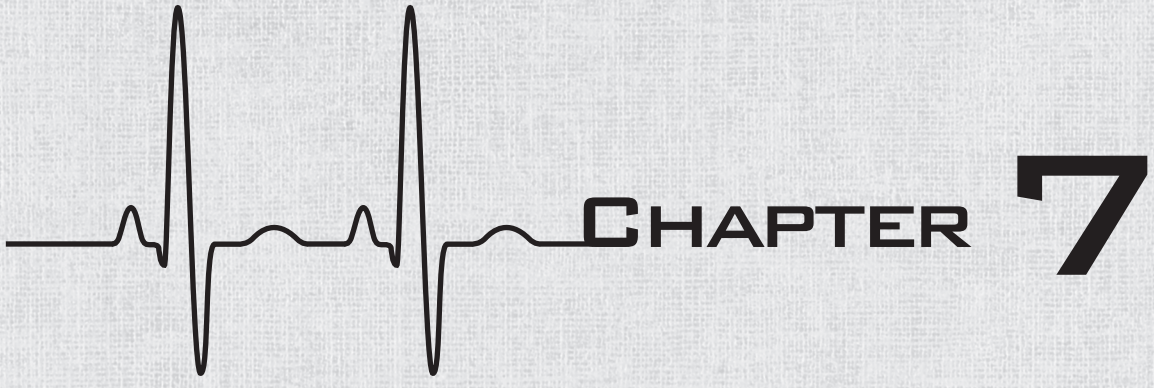
Validation Of Not-In-Trial Simulation:

Predicting Cardiovascular Risk From Clinical Trial Data Using Cisapride

ASY Chain (1, 2), JP Dieleman (2), M Danhof (1),
MCJM Sturkenboom (2, 3), OE Della Pasqua (1,4)

(1) Leiden/Amsterdam Centre for Drug Research,
Division of Pharmacology, Leiden University, Leiden, The Netherlands
(2) Department of Medical Informatics and Epidemiology,
Erasmus Medical Centre, The Netherlands
(3) Department of Epidemiology, Erasmus Medical Centre, The Netherlands
(4) Clinical Pharmacology and Discovery Medicine,
GlaxoSmithKline, Stockley Park, Uxbridge, United Kingdom

Submitted for publication



Chapter 7

ABSTRACT

Previously, a tool was developed to translate clinical trial findings to real life settings by resolving the discrepancies between pre- vs. post-market estimates of QTc-interval prolongation using d,l-sotalol as the paradigm compound. The tool consisted of a prediction model for QTc measurements based on baseline values, QTc model for drug effect from clinical population and effects of concomitant drugs and comorbidities. The objective of this study was to apply the same model to cisapride, as a form of external model validation. There were 87 new cisapride users in the Rotterdam Study cohort (60 female and 27 male) with the average observed QTc of 432.4 msec and 423.5 msec *for women and men respectively. The average simulated QTc for female was 435.8 (sd. 15.83) and 424.1 (sd. 19.5) for male. It was verified that same tool can be used to predict real life QTc values for cisapride users gaining confidence in the use of Not-In-Trial Simulation as a risk mitigation tool.*

7.1 INTRODUCTION

One of the main objectives of clinical trials is to test whether a new chemical entity (NCE) is efficacious towards the condition it is targeted to treat and whether it is safe for human usage. Together, evidence on efficacy and safety allow for the assessment of the expected risk-benefit ratio [1]. However, such estimates are only readily available for events which occur frequently enough, i.e. those with high incidence. Less frequent or rare events require larger populations or longer trials and are usually evaluated during the post-approval phase [2]. 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; QTc-interval prolongation is an example of such a case.

QTc-interval has been considered a risk factor for torsade de pointes (TdP), and has become a major cause for drug withdrawal [3]. Mitigation measures have been implemented by the International Conference on Harmonisation (ICH) [4], which proposes the use of thorough-QT (TQT) studies to detect QTc-interval prolongation as part of the clinical evaluation package for regulatory submissions. Despite these efforts, important limitations are being overlooked because of the mandatory status these studies have. In addition to the lack of quantitative evidence, i.e., the absence of explicit details on the underlying concentration-QTc effect relationship, the current approach often relies on data from healthy subjects who are enrolled according to strict inclusion / exclusion criteria, which does not necessarily reflect the demographic characteristics in the patient population. Moreover, in a TQT trial, supra-therapeutic levels of the drug are used under the assumption that the resulting effects would be representative of the risk in patients, who may be more vulnerable or sensitive to the pro-arrhythmic effects.

Nonetheless, mounting evidence suggests that current TQT protocols may provide biased estimates of the actual risk of QTc-interval prolongation, with high rates of false positive results due to the accepted data analysis method [5]. These findings raise important questions both from a drug development and from a pharmacovigilance perspective. It is in everyone's best interest to ensure accurate assessment of the risk-benefit ratio, so that decision making

regarding the development, approval or risk management is effective and timely.

Recently, to address this issue we have developed a model-based approach that allows better translation of clinical trial findings into real life conditions. Thanks to the use of a Bayesian hierarchical model, it is possible to integrate pharmacokinetic-pharmacodynamic relationships with disease and other patient-specific characteristics, discriminating drug-induced effects under controlled and uncontrolled conditions [6]. In fact, we have shown that the use of model parameter estimates in conjunction with *Not-In-Trial-Simulations* (NITS) provides a framework for evaluating drug vs. disease-specific properties, which facilitates the assessment of causality.

Unfortunately, those who defend the *status quo*, i.e., the requirement for a TQT study as the best approach to gather evidence on the propensity for QTc-interval prolongation, continue to use patient safety as a cogent argument for adhering to current practices. There is an urgent need to consider alternative methods which do not only ensure lower false positive rates, but also provide accurate estimation of the pro-arrhythmic risk taking into account other factors relevant to the target population in clinical practice. Essentially, there needs to be an increase in operational effectiveness in drug development. We envisage therefore the use of a Not-In-Trial Simulation as a risk management tool for the evaluation of causality and assessment of the impact of mitigation measures as shown in Figure 1.

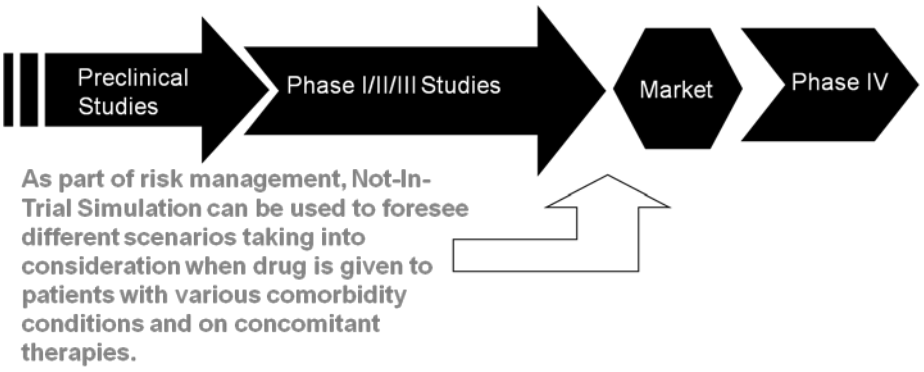


Figure 1. Not-In-Trial Simulation can be used as a framework for risk management and assessment of the impact of mitigation measures. Most importantly, it provides quantitative estimates of the contribution of drug-related vs. drug-unrelated factors.

In a previous publication we have shown that the method resolves the discrepancies in pre- vs. post-market estimates of QTc-interval prolongation. *d,l*-Sotalol was used as the paradigm compound in which QTc changes are described as a function of baseline QTc measurements, drug-induced effects, and contributions of concomitant drugs and comorbidities. Furthermore, the model also demonstrated that the relationship between exposure and QTc-interval is best characterised in terms of drug concentrations rather than the traditional concept of defined daily dose (DDD) in epidemiology studies. The objective of the current investigation is therefore to demonstrate the generalisability of the approach to other areas of drug development by performing an external validation. External validation is an essential step for the use of predictive or prognostic tools in clinical practice [7].

Here we apply hierarchical modeling and Not-In-Trial Simulations to characterise the impact of pro-arrhythmic effects of cisapride in real life conditions. Cisapride is a gastroprokinetic agent which is used to increase motility in the upper gastrointestinal tract. It was first introduced in 1990 after showing promising clinical trial results in healing and producing symptomatic improvement among patients with esophagitis [8-12] as well as improving symptoms among those with symptomatic heartburn [13, 14], preventing the provocation of heartburn by a fatty meal in susceptible individuals [15] and reducing relapse rates among those with healed esophagitis [16].

From a safety perspective, evidence including post-marketing data from two study groups totaling over 23 000 patients suggested that the drug was safe, with diarrhea, abdominal pain, nausea or vomiting, headache and constipation being the most frequent reported adverse events [17]. Furthermore, serious cardiac rhythm disorders were not found to be associated with cisapride use in a review of almost 37 000 patients who were prescribed cisapride in the UK and Canada [18]. However, reports of cardiac events began to accumulate, beginning with a report of palpitations [19], followed by culmulating instances of unusual tachyarrhythmia, torsades de pointes, ventricular fibrillation and sudden death [20]. Eventually, it came to be recognised that cisapride use could result in the prolongation of the QT-interval and, thereby increasing the risk of arrhythmia. It was later revealed that cisapride's arrhythmogenic potential existed at the cellular and molecular level. In *in vitro* studies, cisapride was found to be a potent and dose-dependent blocker of the human ether-a-go-go-related gene (hERG)

channel [21], which is the main channel responsible for the repolarisation phase of the cardiac action potential [22].

In 1995, a black box warning contraindicating the use of cisapride among those taking drugs that affected its metabolism was issued by the FDA; at that time thirty-four cases of torsades de pointes, twenty-three of prolonged QT-interval and four deaths had been reported (103). A further warning was issued in 1996 regarding to its use with medications that also prolonged the QT-interval and in patients with conditions that predisposed to cardiac arrhythmias. The black box warning was again expanded in 1998. However, the subsequent realisation that serious adverse cardiac events could occur even among low-risk groups, including children, coupled with the documentation of continued cisapride use in contraindicated situations (112) led to the withdrawal of the drug worldwide in July 2000. Thus cisapride is an example, given its sensational rise and demise, where a tool should be in place that can anticipate events, such as QT-interval prolongation, in the future for new compounds in clinical development.

7.2 METHODS

Evaluation of the effect of age on QTc-interval

Baseline information from healthy subjects participating in seven Phase I studies was used for the development of the relationship between baseline QTc-interval and age. The meta-data consisted of approximately 1000 subjects. These data were combined with elderly patient data that were classified as healthy in the Rotterdam Study cohort described below. The healthy elderly subjects were defined as those who were not on any medications and not have any comorbidity conditions recorded at the time of the ECG measurement. The Bazett formula was used as the RR correction factor for QT-intervals as it was a consistent output for all the data used in the study. Only QTc values vs. age relationship was explored for this analysis while the effect of age on RR-interval was not explored.

In the original simulation model, a quadratic regression was used to describe the age-dependent QTc data. Based on the resulting parameter estimates, the regression was reduced to a linear relationship in R2.8 (The R project -

http://www.r-project.org/), where QTc = 0.5096(Age) + 386.52 for the male population and QTc = 0.8617(Age) + 371.34 for the female population.

ECG measurement in clinical trials and real life population

A 10 sec 12-lead ECG, resulting on average 8-10 heart beats, was recorded with an ACTA electrocardiograph (ESAOTE, Florence, Italy) at a sampling frequency of 500 Hz and stored digitally. All ECGs were processed by the Modular ECG Analysis System (MEANS) to obtain ECG measurements. The MEANS program has been evaluated extensively and reported elsewhere [28-30]. MEANS determines common onsets and offsets for all 12 leads together for one representative averaged beat, with the use of template matching techniques [29], until the end of the T wave. To adjust for heart rate, Bazett's formula (QTc = QT / RR ^ ½) was used [31].

Evaluation of drug-induced effects in healthy subjects

Data on the pharmacokinetics and pharmacodynamics of cisapride was derived from a randomised, placebo-controlled, dose-escalating Phase-I trial with up to 6 sessions. The primary objective was to assess the cardiovascular safety of single doses of 10, 20, 40, 80, 130 and 200 mg cisapride. Secondary objectives were to determine the pharmacokinetics of cisapride at these single doses, to search for a correlation between pharmacokinetics and cardiovascular parameters, and to evaluate an experimental beat-to-beat computerised analysis of electrocardiograms (ECGs) recorded by Holter monitoring in comparison with the analysis of traditional 12-lead timed ECGs taken at pharmacokinetic blood sampling times. Subjects were randomised in 4 blocks to treatments so that for each block and for each dose, 5 subjects received cisapride and 1 subject received placebo. There was a 48-hour wash-out interval between consecutive doses. A total of 24 healthy subjects participated in the trial. The dose escalation was stopped after dosing with 130 mg cisapride (or placebo) on Day 9, because of the investigator's medical judgment of a potential risk if dose escalation would proceed to the highest dose level of 200 mg of cisapride.

Other relevant eligibility criteria for inclusion into the study for the purposes of our investigation were non-smoking status and body mass index between 18-28 kg/m^2, inclusive. Subjects with a history of syncope and a family history of sudden death or long QT syndrome were excluded at screening. In

addition, subjects showing an increase in number of premature ventricular beats at a heart rate of >120 bpm, >500 premature ventricular beats per 24 hours, abnormal STU dynamics, a RR-interval > 3 sec, an average heart rate < 50 bpm or > 95 bpm, apparent QT prolongation, ventricular tachycardia of >3 beats at a rate of 100bpm, QRS-interval > 0.12 sec, apparent ST abnormalities, second degree atrioventricular block, supraventricular tachycardias with rates > 150 bpm, QTc-intervals above 450 or 430 msec (for male and female, respectively) were also excluded.

Modeling of QTc-interval prolongation

The pharmacokinetic model (PK) as well as the pharmacokinetic pharmacodynamic (PKPD) model of cisapride was taken as reported from the analysis report for the study to demonstrate that Not-In-Trial Simulations can be applied to the current workflow. In the report, it was indicated that the models were fitted to the data using NONMEM V software version 1.1, by applying the first-order conditional estimation (FOCE) method. The package was installed on a PC platform using Compaq Visual Fortran version 6.6 under Microsoft Windows 2000.

A two-step sequential approach was adopted, including a pharmacokinetic modeling step followed by a pharmacodynamic modeling step. The pharmacokinetic model included a two-compartment disposition and a sequential zero- and first-order absorption with a lag time. The relationship between cisapride concentration in plasma and heart rate, RR-interval, uncorrected QT-intervals, corrected QT-intervals using the Bazett and Fridericia formulas as well as other correction formulas, was explored graphically, and by means of linear regression.

The ECG parameters were fitted linearly on the cisapride concentrations. The regressions were based on the ECGs taken immediately before the corresponding pharmacokinetic blood sampling. Plasma concentrations that were below the lower limit of quantification (5 ng/mL) were replaced by zero. Since both heart rate and QT-interval tended to increase with higher cisapride concentrations, all QT corrections also increased with concentration. The tests for the model parameters were always statistically significant ($p=0.0001$), indicating that the relationship between ECG parameters and cisapride concentrations was most likely not an observation by chance.

Drug-induced QTc-intervals in patients was simulated assuming full compliance to treatment. Population parameter estimates were used for the simulations (slope = 0.179 msec/ng/ml, and intercept (QT_0) = 393.5 msec).

Evaluation of the effect of intrinsic and extrinsic factors on QTc-interval in real life patients

Data from a real life population setting was obtained from the Rotterdam Study. The study is a prospective population-based cohort study, which started with a baseline visit between 1990 and 1993. All inhabitants of a suburb in Rotterdam, Ommoord, aged 55 years and over were invited to participate. Of the 10 275 subjects invited, 7983 (78%) gave their written informed consent and took part in the baseline examination. The study was approved by the Medical Ethics Committee of the Erasmus Medical Center, Rotterdam, The Netherlands. Objectives and methods of investigation have been described in detail in other publications [32]. All participants were visited at home for a standardised questionnaire and 7151 were subsequently examined at the research center. A second, third and fourth follow up visit took place, respectively, between 1993 - 1995, 1997 - 1999 and 2002 - 2004. In addition to follow-up examinations, the cohort is continuously monitored for major morbidity and mortality through linkage with general practitioner and municipality records.

All subjects who started treatment with cisapride during follow-up were included in the present study, with the exception of prevalent users and subjects with left ventricular hypertrophy (LVH), right bundle-branch block (RBBB) and left bundle-branch block (LBBB), who were excluded. The reason for the exclusion of these patients is that LVH, RBBB and LBBB can cause secondary repolarisation changes and atrial fibrillation can cause difficulties in measuring QT-intervals. From the remaining population, "new" cisapride users were identified as those receiving their first prescription after enrolment into the study. Our analysis included all QTc measurements after the start of cisapride for all new users.

Comorbidities were screened and identified according to standard clinical criteria. Hypertension was defined as systolic blood pressure > 160 mm Hg or diastolic blood pressure > 100 mm Hg and/or use of antihypertensive medications, encompassing grade 2 and grade 3 hypertension compounds according to the World Health Organisation (WHO) criteria [33]. Diabetes

mellitus was defined as the use of blood glucose-lowering medication and / or a non-fasting or post-load serum glucose level of 11.1 nmol/L or higher according to the WHO [34]. A history of myocardial infarction was assessed by self-report checked with records from the general practitioner or cardiologist and/or electrocardiographic evidence. All reported myocardial infarctions were verified against the medical records as described in detail previously [35]. Assessment of heart failure at baseline and during follow-up was assessed by reviewing all medical records for the occurrence of at least two signs and symptoms suggestive of heart failure or the use of medication for the indication heart failure and hospital discharge letters. Cases of incident heart failure were obtained by continuously monitoring the participants [36, 37]. The ankle arm index (AAI) was used as potential predictor of cardiovascular diseases and mortality [38].

The impact of co-medications and comorbidities were evaluated using a linear regression where the absolute differences in QTc prolongation between taking *d,l*-sotalol alone and in conjunction with co-medications and comorbidities were calculated with standard deviation. In the original model, the increases of QTc-interval values were derived from non-*d,l*-sotalol users. Therefore, the same values can be applied here for the validation of the simulation model using cisapride. These values are indicated in table 1.

	Increase in QTc [msec]	
	Male (CI)	**Female (CI)**
Heart Failure	7.7 (+/- 4.02)	5.8 (+/- 3.49)
Myocardial Infarction	3 (+/- 2.11)	3.2 (+/- 2.31)
Arrhythmia	18.5 (+/- 9.24)	20.4 (+/- 9.51)
Diabetes	3.0 (+/- 2.28)	5.4 (+/- 2.0)
Hypertension	0.117 (+/- 0.08)	-0.142 (+/- 0.05)

*Table 1. Increase in QTc [*msec*] due to various concomitant medications and comorbidity conditions*

Not-In-Trial Simulation

The final distribution of QTc values associated with all causal factors was simulated stepwise in R2.8. Differences between observed and expected QTc-intervals were compared non-parametrically by visual inspection.

7.3 RESULTS

PKPD modeling of QTc-interval prolongation

Drug effect-concentration relationships are routinely investigated as part of the drug development process. The relationship of cisapride, shown in Figure 2, has been directly taken from the original Phase I clinical trial report. This illustrate that Not-In-Trial Simulation can be applied in addition to the current workflow, rather than changing the drug development paradigm.

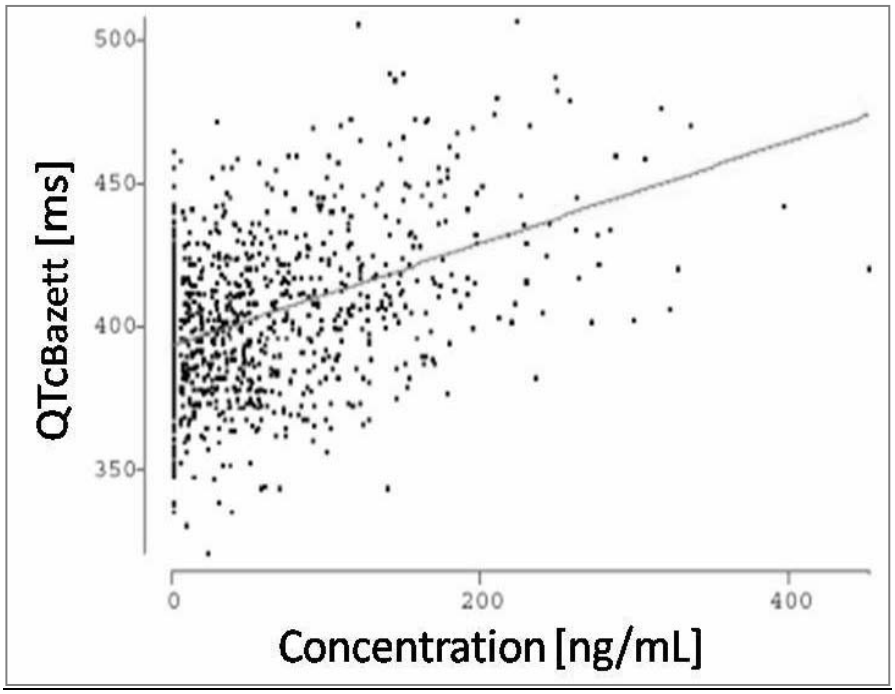
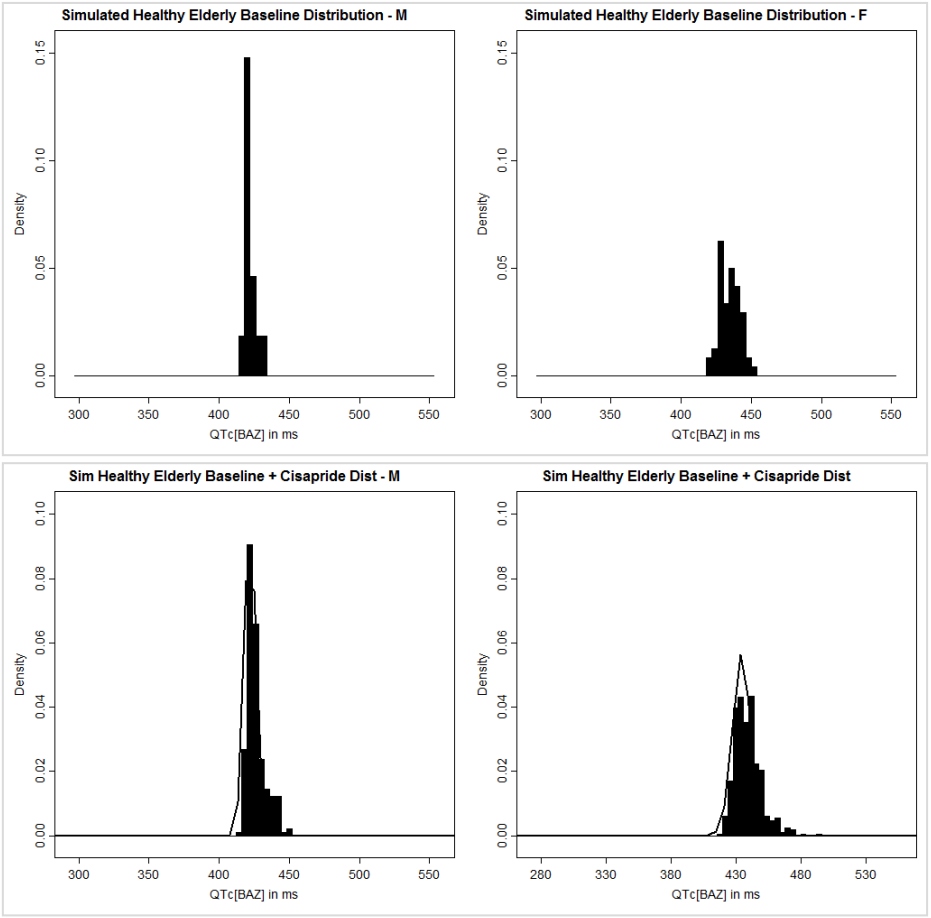


Figure 2. Drug effect – concentration relationship of cispride.

Not-In-Trial Simulation

By combining the drug model with the age-dependent baseline relationship, along with the added contributions from comorbidities and concomitant medications, the progression of QTc-interval prolongation in the real life cisapride users is displayed in Figure 3. The distribution of the observed QTc

values (light colours) fell within the simulated distribution (dark colours). After applying the model developed using *d,l*-sotalol, it was verified that same tool can be used to predict real life QTc values for cisapride users.



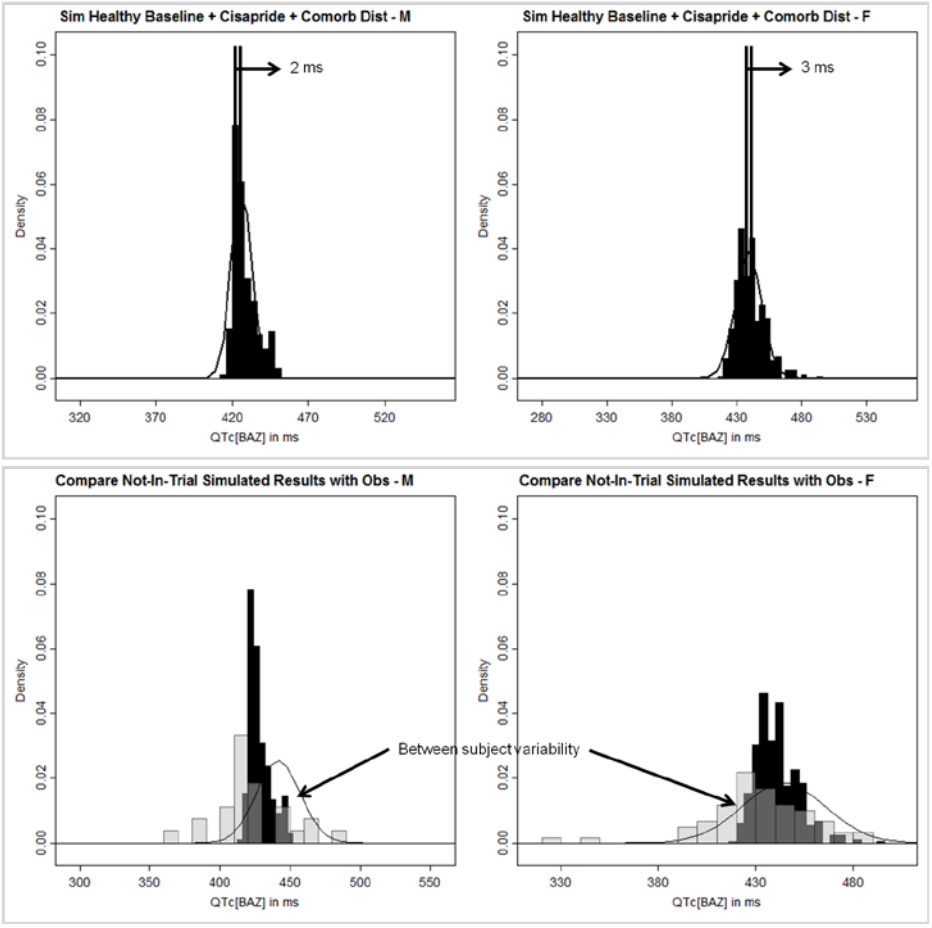


Figure 3. Not-In-Trial Simulation of real life cispride users. From hypothetical age-dependent baseline values, to drug-induced QTc prolongation, to added effects from comorbidity conditions and co-medication and finally comparing the simulated values to observed measurements with between subject variability (BSV) shown.

7.4 DISCUSSIONS

Operational effectiveness

Understanding the actual benefit-risk ratio in the target population has become paramount for the development of successful candidate molecules and prevention of attrition at the post-approval phase. The assessment of the

propensity for QTc-interval prolongation is a very pertinent example of this requirement, as these events have become a major cause of drug withdrawal.

Current regulatory and clinical efforts regarding the effective use of drugs 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 [1]. 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 [23-25]. 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 [26, 27].

In the current investigation we have shown once more that evidence synthesis can indeed be performed prospectively by the implementation of model-based approaches in which both fixed effects (drug-related as well as disease-related factors) as well random effects (including parameter and model uncertainty) are integrated. Our results reveal that Not-In-Trial Simulations can be used to predict real life impact of drug intervention on actual patients under real life conditions. QTc measurements in cisapride users were accurately described especially the extremely high values. The lower values were not as well described. One possible explanation that cannot be excluded is that those figures reflect measurement errors, despite quality control and review of the data entry. In fact, the readings showed values that are considerably low for an elderly population. Another possible explanation is due to the pharmacological effects of cisapride where short term usage would induce QTc prolongation; however, longer therapy might also cause a shortening phenomenon which cannot be anticipated by the model.

Validation procedures

Various requirements exist for the assessment of the predictive or prognostic value of a model. Predictive value refers to the ability to assess the probability of a target outcome while prognostic value defines the factors that predict prevalence. Among other things, internal and external validation procedures provide evidence of consistency, accuracy and precision of the estimates.

It is also essential to consider specificity and sensitivity of the model in discriminating drug-induced effects from other confounding factors, in particular when drug-drug interactions or drug-disease interactions may occur which are compound-specific and may not have been identified yet, i.e., where uncertainty exists about the actual effects.

The major limitation to our approach at this point of development is the exclusion of random effects. The inclusion of random effects and variabilities would further improve the assessment and predictability of the exposure-effect relationship. Further work will be required to incorporate various sources of variability and uncertainties in the model. Specifically, hierarchical modeling techniques can be applied to the relationship describing age and QTc values in healthy subjects. Alternatively, since age, concomitant medications as well as comorbidity conditions can all be considered systems-related factors affecting QT/QTc-intervals, a combined estimation approach can be taken with uncertainties built in.

In summary, having been able to successfully predict QTc values in real life for two different compounds using an age-dependent baseline QTc relationship, in conjunction with the effects from co-medications and comorbidities, we have gained confidence that Not-In-Trial Simulation will become a framework for signal detection in clinical development. Although our work has been limited to QTc-interval prolongation, we envision the application of the concept for any adverse event. This framework offers advantages compared to meta-analysis where the resulting model can only describe a large, pooled dataset rather than being predictive. Based on Not-In-Trial Simulation scenarios, investigators will have the freedom to explore hypothetical populations and test different doses which are not confined to the observed data.

REFERENCES

[1] Curtin F, Schulz P. Assessing the benefit: risk ratio of a drug--randomized and naturalistic evidence. Dialogues Clin Neurosci 2011;13(2):183-90.

[2] Benefit-Risk Balance for Marketed Drugs: Evaluating Safety Signals. In: IV CWG, editor. Geneva, Switzerland; 2004.

[3] Shah RR. Drug-induced QT interval prolongation: regulatory perspectives and drug development. Ann Med 2004;36 Suppl 1:47-52.

[4] Guidance for Industry: ICH E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrythmic Potential for Non-Antiarrythmic Drugs. U.S. Depart of Health and Human Services. Food and Drug Administration. (2005)

[5] Hutmacher MM, Chapel S, Agin MA, Fleishaker JC, Lalonde RL. Performance characteristics for some typical QT study designs under the ICH E-14 guidance. J Clin Pharmacol 2008 Feb;48(2):215-24.

[6] Chain AS, Krudys KM, Danhof M, Della Pasqua O. Assessing the Probability of Drug-Induced QTc-Interval Prolongation During Clinical Drug Development. Clin Pharmacol Ther 2011 Dec;90(6):867-75.

[7] Bleeker SE, Moll HA, Steyerberg EW, Donders AR, Derksen-Lubsen G, Grobbee DE, et al. External validation is necessary in prediction research: a clinical example. J Clin Epidemiol 2003 Sep;56(9):826-32.

[8] Arvanitakis C, Nikopoulos A, Theoharidis A, Giannoulis E, Vagios I, Anthopoulou H, et al. Cisapride and ranitidine in the treatment of gastro-oesophageal reflux disease--a comparative randomized double-blind trial. Aliment Pharmacol Ther 1993 Dec;7(6):635-41.

[9] Baldi F, Bianchi Porro G, Dobrilla G, Iascone C, Lobello R, Marzio L, et al. Cisapride versus placebo in reflux esophagitis. A multicenter double-blind trial. J Clin Gastroenterol 1988 Dec;10(6):614-8.

[10] Galmiche JP, Fraitag B, Filoche B, Evreux M, Vitaux J, Zeitoun P, et al. Double-blind comparison of cisapride and cimetidine in treatment of reflux esophagitis. Dig Dis Sci 1990 May;35(5):649-55.

[11] Geldof H, Hazelhoff B, Otten MH. Two different dose regimens of cisapride in the treatment of reflux oesophagitis: a double-blind comparison with ranitidine. Aliment Pharmacol Ther 1993 Aug;7(4):409-15.

[12] Schutze K, Bigard MA, Van Waes L, Hinojosa J, Bedogni G, Hentschel E. Comparison of two dosing regimens of cisapride in the treatment of reflux oesophagitis. Aliment Pharmacol Ther 1997 Jun;11(3):497-503.

[13] Castell DO, Sigmund C, Jr., Patterson D, Lambert R, Hasner D, Clyde C, et al. Cisapride 20 mg b.i.d. provides symptomatic relief of heartburn and related symptoms of chronic mild to moderate gastroesophageal reflux disease. CIS-USA-52 Investigator Group. Am J Gastroenterol 1998 Apr;93(4):547-52.

[14] Richter JE, Long JF. Cisapride for gastroesophageal reflux disease: a placebo-controlled, double-blind study. Am J Gastroenterol 1995 Mar;90(3):423-30.

[15] Castell D, Silvers D, Littlejohn T, Orr W, Napolitano J, Oleka N, et al. Cisapride 20 mg b.d. for preventing symptoms of GERD induced by a provocative meal. The CIS-USA-89 Study Group. Aliment Pharmacol Ther 1999 Jun;13(6):787-94.

[16] Toussaint J, Gossuin A, Deruyttere M, Huble F, Devis G. Healing and prevention of relapse of reflux oesophagitis by cisapride. Gut 1991 Nov;32(11):1280-5.

[17] Wager E, Tooley PJ, Pearce GL, Wilton LV, Mann RD. A comparison of two cohort studies evaluating the safety of cisapride: Prescription-Event Monitoring and a large phase IV study. Eur J Clin Pharmacol 1997;52(2):87-94.

[18] Walker AM, Szneke P, Weatherby LB, Dicker LW, Lanza LL, Loughlin JE, et al. The risk of serious cardiac arrhythmias among cisapride users in the United Kingdom and Canada. Am J Med 1999 Oct;107(4):356-62.

[19] Olsson S, Edwards IR. Tachycardia during cisapride treatment. BMJ 1992 Sep 26;305(6856):748-9.

[20] Smalley W, Shatin D, Wysowski DK, Gurwitz J, Andrade SE, Goodman M, et al. Contraindicated use of cisapride: impact of food and drug administration regulatory action. JAMA 2000 Dec 20;284(23):3036-9.

[21] Mohammad S, Zhou Z, Gong Q, January CT. Blockage of the HERG human cardiac K+ channel by the gastrointestinal prokinetic agent cisapride. Am J Physiol 1997 Nov;273(5 Pt 2):H2534-8.

[22] Drolet B, Khalifa M, Daleau P, Hamelin BA, Turgeon J. Block of the rapid component of the delayed rectifier potassium current by the prokinetic agent cisapride underlies drug-related lengthening of the QT interval. Circulation 1998 Jan 20;97(2):204-10.

[23] Lexchin J. Progressive Licensing of Drugs: Music or Noise? Healthc Policy 2008 May;3(4):11-20.

[24] Yeates N, Lee DK, Maher M. Health Canada's Progressive Licensing Framework. CMAJ 2007 Jun 19;176(13):1845-7.

[25] Hebert PC, Stanbrook M, Sibbald B, Flegel K, MacDonald N, Attaran A. Progressive licensing needs progressive open debate. CMAJ 2007 Jun 19;176(13):1801, 3.

[26] McKenna C, Claxton K. Addressing adoption and research design decisions simultaneously: the role of value of sample information analysis. Med Decis Making 2011 Nov-Dec;31(6):853-65.

[27] Presanis AM, De Angelis D, Goubar A, Gill ON, Ades AE. Bayesian evidence synthesis for a transmission dynamic model for HIV among men who have sex with men. Biostatistics 2011 Oct;12(4):666-81.

[28] de Bruyne MC, Kors JA, Hoes AW, Kruijssen DA, Deckers JW, Grosfeld M, et al. Diagnostic interpretation of electrocardiograms in population-based research: computer program research physicians, or cardiologists? J Clin Epidemiol 1997 Aug;50(8):947-52.

[29] van Bemmel JH, Kors JA, van Herpen G. Methodology of the modular ECG analysis system MEANS. Methods Inf Med 1990 Sep;29(4):346-53.

[30] Willems JL, Abreu-Lima C, Arnaud P, van Bemmel JH, Brohet C, Degani R, et al. The diagnostic performance of computer programs for the interpretation of electrocardiograms. N Engl J Med 1991 Dec 19;325(25):1767-73.

[31] Bazett HC. The time relations of the blood-pressure changes after excision of the adrenal glands, with some observations on blood volume changes. J Physiol 1920 Feb 20;53(5):320-39.

[32] Hofman A, Grobbee DE, de Jong PT, van den Ouweland FA. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. Eur J Epidemiol 1991 Jul;7(4):403-22.

[33] 1999 World Health Organization-International Society of Hypertension Guidelines for the Management of Hypertension. Guidelines Subcommittee. J Hypertens 1999 Feb;17(2):151-83.

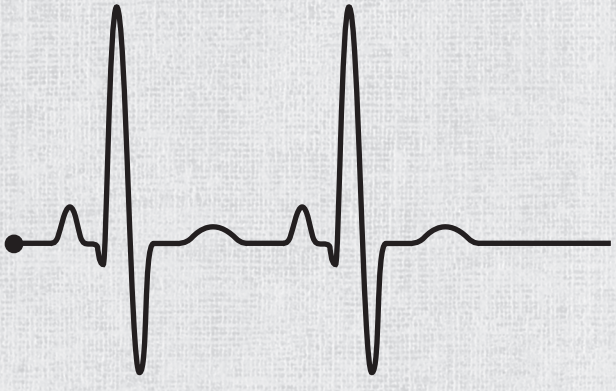
[34] Diabetes Mellitus. Technical Reports series 894. Geneva; 1985.

[35] Bots ML, Hoes AW, Koudstaal PJ, Hofman A, Grobbee DE. Common carotid intima-media thickness and risk of stroke and myocardial infarction: the Rotterdam Study. Circulation 1997 Sep 2;96(5):1432-7.

[36] Bleumink GS, Knetsch AM, Sturkenboom MC, Straus SM, Hofman A, Deckers JW, et al. Quantifying the heart failure epidemic: prevalence, incidence rate, lifetime risk and prognosis of heart failure The Rotterdam Study. Eur Heart J 2004 Sep;25(18):1614-9.

[37] Mosterd A, Hoes AW, de Bruyne MC, Deckers JW, Linker DT, Hofman A, et al. Prevalence of heart failure and left ventricular dysfunction in the general population; The Rotterdam Study. Eur Heart J 1999 Mar;20(6):447-55.

[38] International classification of diseases and health related problems. Geneva; 1992.

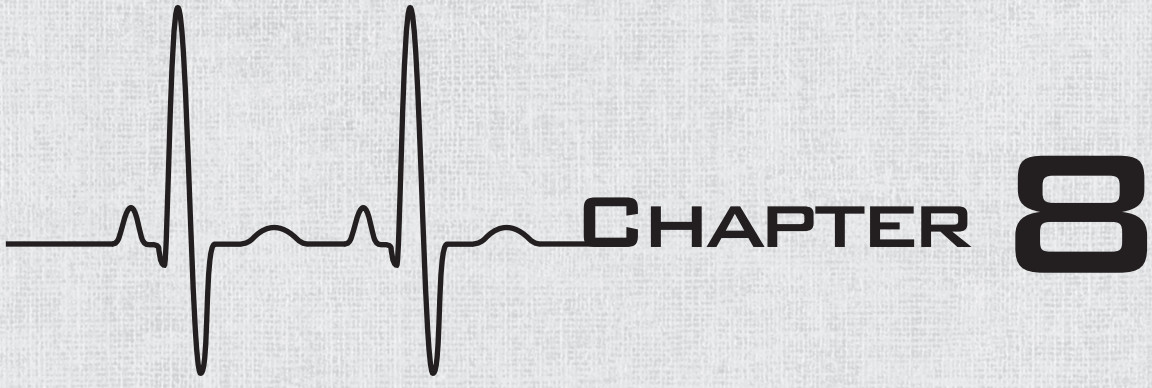


Can QTc Prolongation >10 msec Predict Sudden Cardiac Death?

Anne SY Chain (1), Jeanne Dieleman (2),
Meindert Danhof (1),Miriam CJM Sturkenboom (2)
Oscar E Della Pasqua (1, 3)

(1) Division of Pharmacology, Leiden/Amsterdam
Center for Drug Research, Leiden University, The Netherlands
(2) Department of Medical Informatics,
Erasmus University Medical Center, The Netherlands
(3) GlaxoSmithKline, Clinical Pharmacology/Modelling
and Simulation, United Kingdom

Submitted for publication



Chapter 8

ABSTRACT

Context: Currently, specific clinical studies are mandated before drug approval to ensure accurate evaluation of the risk of drug-induced arrhythmias. The underlying assumption for this requirement 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.

Objective: To challenge the notion of increased TdP risk with QTc-interval prolongation ≥10msec and to explore whether sudden cardiac death (SCD), a potential surrogate for TdP, is correlated with increases from baseline QTc or to high baseline values.

Design, Setting, and Patients: The incidence of SCD in the population was derived from data in the Rotterdam Study cohort observed between 1997 and 2004 (n=282).

Main Outcome Measures: Trend-analysis was applied to evaluate SCD rates in patients showing increasing, decreasing and no change in QTc-interval measurements. A sensitivity analysis was also performed on the impact of the time span between death and last ECG measurement.

Results: Incidence rates of SCD increased significantly with higher baseline QTc values. This increase was independent of the time elapsed between the last ECG and death. In contrast to current belief, the incidence rates in subjects showing QTc prolongation <10 msec and ≥10 msec varied from 4.6 to 4.0 cases per 1000 person years (p<0.261).

Conclusions: QTc-interval prolongation by ≥10 msec does not appear to affect the incidence rate of SCD. Instead, increased rates were found to better correlate with high baseline QTc values. Despite the limited size of the cohort, these findings prompt for careful revision of the guidelines for evaluating the pro-arrhythmic risk of new medicines.

8.1 INTRODUCTION

Approximately 50% of all cardiac deaths are estimated to be of sudden nature.[1-3] Sudden cardiac death (SCD) is defined as "natural death due to cardiac causes, heralded by abrupt loss of consciousness within one hour after the onset of acute symptoms; pre-existing heart disease may have been known to be present, but the time and mode of death are unexpected".[4] Irrespective of other triggering conditions, ventricular arrhythmias associated with an acquired long-QT syndrome (LQTS) are considered a risk factor for SCD.[3] In addition, it is also assumed that QT-interval prolongation can provoke torsades de pointes (TdP), which in turn can either resolve spontaneously or deteriorate into ventricular fibrillation and potentially into SCD.[5]

The alleged link between LQTS and SCD has become a real concern for the medical community [3] given that a variety of drugs have been recognised to possess substantial pro-arrhythmic potential by inducing QTc-interval prolongation and consequently increasing the risk of TdP. Whilst the mechanism by which prolonged QTc-intervals can degrade into fatal arrhythmias is well described in various publications [6,7], there is still an ongoing debate in the literature on the clinical significance of its prolongation [1]. Yet, QTc-interval prolongation has become one of the most common causes of post-market withdrawal as well as restriction of use of drugs.[7]

Several large population-based studies evaluating 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[8-14]. Only few of these studies specifically explored events such as sudden cardiac death and when they did, the number of cases was often too small to detect statistically significant differences.[9,15,16]. Nevertheless, SCD has been found to be associated with high values of QTc-interval in a study conducted by Straus *et al.*.[17] In view of the potential implications of drug-induced LQTS, the use of QTc-interval prolongation has been proposed as a marker of cardiovascular risk for both pre- and post-market studies, regardless of the supporting evidence.[18]

To mitigate such risks, the European Agency for Evaluation of Medicinal Products (EMA) was the first to issue a "points to consider" document from the Committee for Proprietary Medicinal Products (CPMP) in 1997.[19] In the early 2000s, regulatory scrutiny continued to increase [20], leading to the development of the latest guideline from the International Conference on Harmonisation (ICH), topic E14, which introduced the requirement for a thorough-QT (TQT) study. A TQT study will only be considered negative for a new chemical entity (NCE) if the 95% upper interval of the largest time-matched QTc measurements after active treatment does not include 10 msec after adjusting for baseline and placebo effects.[21] Yet, the guideline states that it is not clear whether arrhythmia development is more closely related to an increase in the absolute QT or QTc.

Irrespective of the premises adopted by regulatory agencies and recent advancements in the detection of QT-interval prolongation, it remains unclear whether TdP and SCD are indeed correlated with each other in a cause-and-effect manner. As a result, 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. Among other things, it is of interest to investigate whether increases in QTc values, calculated as changes from baseline measurements or absolute measured QTc values, are associated with SCD using traditional epidemiological methodologies. The aim of this study was therefore to explore whether the incidence rate of SCD is correlated with high baseline values or increases from baseline QTc-interval.

8.2 METHODS

Protocol design and study population

This study was conducted using the Rotterdam Study cohort. All inhabitants of a suburb in Rotterdam, Ommoord, aged 55 years and over (n = 10275), were invited to participate. Of these, 7983 (78%) gave their written informed consent and took part in the baseline examination. Objectives and methods of the Rotterdam Study have been described in detail elsewhere.[22,23] During baseline assessment, all participants were visited at home for a standardised questionnaire and 7151 were subsequently examined at the

research centre. A second, third and fourth follow up visit took place between 1993 and 2004. In addition to follow-up examinations, the cohort was continuously monitored for major morbidity and mortality through linkage with general practitioner and municipality records. Drug prescriptions dispensed to participants have been stored in the database by automated pharmacy procedures. All participants in the Rotterdam Study with at least two consecutive ECG assessments were included in this analysis.

Sudden cardiac death assessment

In case of a fatal event, general practitioners (GP) filled in a questionnaire describing the circumstances of the death, including time since the occurrence of first symptoms, most likely cause of death according to the physician, actual time and place of death. All questionnaires and a copy of the medical records were used to assess if the death could be classified as a sudden cardiac death.[5] If death was witnessed and occurred within one hour after the start of symptoms, we assumed it to be a sudden cardiac death, without additional review of the medical records for a medical history of cardiovascular disease or the presence of cardiovascular risk factors. In case of an un-witnessed death, evidence of cardiac causes was searched for using all available information. Two research physicians coded all reported events independently according to the International Classification of Diseases, 10th edition (ICD-10, sudden cardiac death: I.46). In case of disagreement, consensus was sought. Finally, a cardiologist, whose judgment was considered decisive, reviewed all events.

ECG interpretation and measurements

A 10 sec 12-lead resting ECG (on average 8-10 beats) was recorded with an ACTA electrocardiograph (ESAOTE, Florence, Italy) at a sampling frequency of 500 Hz and stored digitally. All ECGs were processed by the Modular ECG Analysis System (MEANS), which determines the QT-interval from the start of the QRS complex until the end of the T wave. The Bazett's formula (QTc[b]=QT/sqrt(RR)) was used to correct for heart rate differences. Digitally stored ECGs from 6134 and 4415 participants who visited the research centre were available at the time of the first and second visit (86% vs. 70% of the study population), respectively.

The immediate last QTc value on record prior to the case of SCD was used as the measurement associated with death. QTc values were categorised into bins of 10 msec increments as shown in Table 1. For calculating changes in QTc-interval the immediate two last consecutive values on record were used. Two analyses were performed using the differences in QTc (delta QTc). The first one was meant to explore the directionality of the two consecutive measurements: downward (negative delta) where $\Delta QTc<0$, flat (zero delta) with $\Delta QTc=0$, and upward (positive delta) where Δ QTc>0. The second analysis was based on a subsequent dichotomisation of delta QTc values into less than 10 msec and those equal or greater than 10 msec. In all those cases in which measurements could not be obtained or were missing, the patient was excluded from this investigation.

Statistical analysis

The incidence rate of sudden cardiac death was determined by dividing the total number of sudden cardiac death cases by the total number of person-years accumulated in the study population. 95% confidence intervals were calculated based on a Poisson distribution.

Trend analysis of the incidence rates was then performed. SCD rates for delta QTc of <10 msec and $\geq$10 msec were compared with negative, zero and positive changes between consecutive measurements. In addition, a sensitivity analysis was performed to assess parameter precision at different time lapses between death and last ECG measurement at 6, 12, 18, 24, 30 and 36 months.

Binomial proportions testing was performed to determine if the proportions of SCD cases with QTc values above and below the population median QTc value were significant using $p < 0.05$ as a statistical significance level. This consisted of comparing the ratios between the number of deaths associated with QTc-intervals that are below the median QTc value of the entire population *versus* the number of deaths associated with QTc-intervals that are above the population median QTc value. Furthermore, logistic regression calculations were done to determine if predictor variables or regression coefficients in this analysis, namely bins of QTc-intervals, delta QTc of <10 msec or $\geq$ 10msec, and delta QTc of < 0, =0 or >0, were correlated with the probability of occurrence of SCD. A p-value < 0.05 was also chosen as threshold for statistical significance.

QT_C Category	QT_C	Number of Deaths	SCD Incidence Rate / 1000 PY	Confidence Intervals
1	<300	0	0.000	-
2	300-309	0	0.000	-
3	310-319	0	0.000	-
4	320-329	0	0.000	-
5	330-339	0	0.000	-
6	340-349	0	0.000	-
7	350-359	0	0.000	-
8	360-369	3	0.013	0.0026 - 0.0369
9	370-379	2	0.003	0.0004 - 0.0110
10	380-389	5	0.003	0.0001 - 0.007
11	390-399	12	0.003	0.0017 - 0.0057
12	400-409	18	0.003	0.0016 - 0.0044
13	410-419	21	0.002	0.0014 - 0.0035
14	420-429	42	0.004	0.0027 - 0.0051
15	430-439	36	0.004	0.0025 - 0.0049
16	440-449	33	0.004	0.0029 - 0.0058
17	450-459	38	0.007	0.0049 - 0.0095
18	460-469	17	0.005	0.0031 - 0.0086
19	470-479	16	0.009	0.0052 - 0.0147
20	480-489	9	0.010	0.0044 - 0.0182
21	490-499	10	0.016	0.0076 - 0.0890
22	500-509	6	0.014	0.0052 - 0.0309
23	510-519	4	0.014	0.0037 - 0.0348
24	520-529	3	0.025	0.0051 - 0.0727
25	530-539	1	0.009	0.0002 - 0.0485
26	540-549	4	0.085	0.0231 - 0.2168
27	550-559	1	0.064	0.0016 - 0.3549
28	560-569	0	0.000	-
29	>570	1	0.906	0.0057 - 1.259

Table 1. QTc categories and the corresponding QTc values in increment bins of 10msec. The incidence rates are shown for each category with its corresponding confidence intervals.

8.3 RESULTS

Out of 6680 subjects with an ECG measurement, 282 cases of SCD were identified of which 147 cases were witnessed. Only 38 subjects were taking QTc-prolonging medications at the time of the last ECG measurement prior to death. No clear pattern has emerged from exploring the sequences of ECG measurements over time, as shown in the top panel in Figure 1 for seven randomly selected subjects who experienced SCD. The median QTc value associated with all the QTc values included in the analysis is 434 msec. There were 162 SCD cases out of 3235 QTc measurements above the median value, 118 cases out of 3315 measurements below the median value and 2 cases out of 91 measurements at the median value. Based on the binomial proportions test it was concluded that the differences between the two groups were statistically significant ($p = 0.004574$). The overall distributions of all QTc measurements associated with cases of SCD and non SCD show that the fatal event can occur at normal, borderline and prolonged values, as defined by the Committee for Proprietary Medicinal Products (Figure 1, middle and bottom panels). [19]

In contrast, incidence rates did increase with higher baseline QTc values ($p < 0.001$) as shown in Table 1. This increase was consistently observed irrespective of the time elapsed between the last ECG measurement and death (Figure 2). When delta QTc values are compared with regard to its direction of fluctuation, the trend over time is uncorrelated ($p = 0.195$). In Figure 3, negative, zero and positive delta QTc values were compared for 6, 12, 18, 24, 30, and 36 months between last ECG measurement and death. Remarkably, shown in Figure 4, the magnitude of the changes in QTc-interval, < 10 msec or $\geq$ 10 msec, was not correlated to incidence rates of SCD with results varying from 4.6 to 4.0 cases per 1000 person years respectively ($p = 0.261$). Δ QTc values were also compared for 6, 12, 18, 24, 30, and 36 months between last ECG measurement and death with insignificant results.

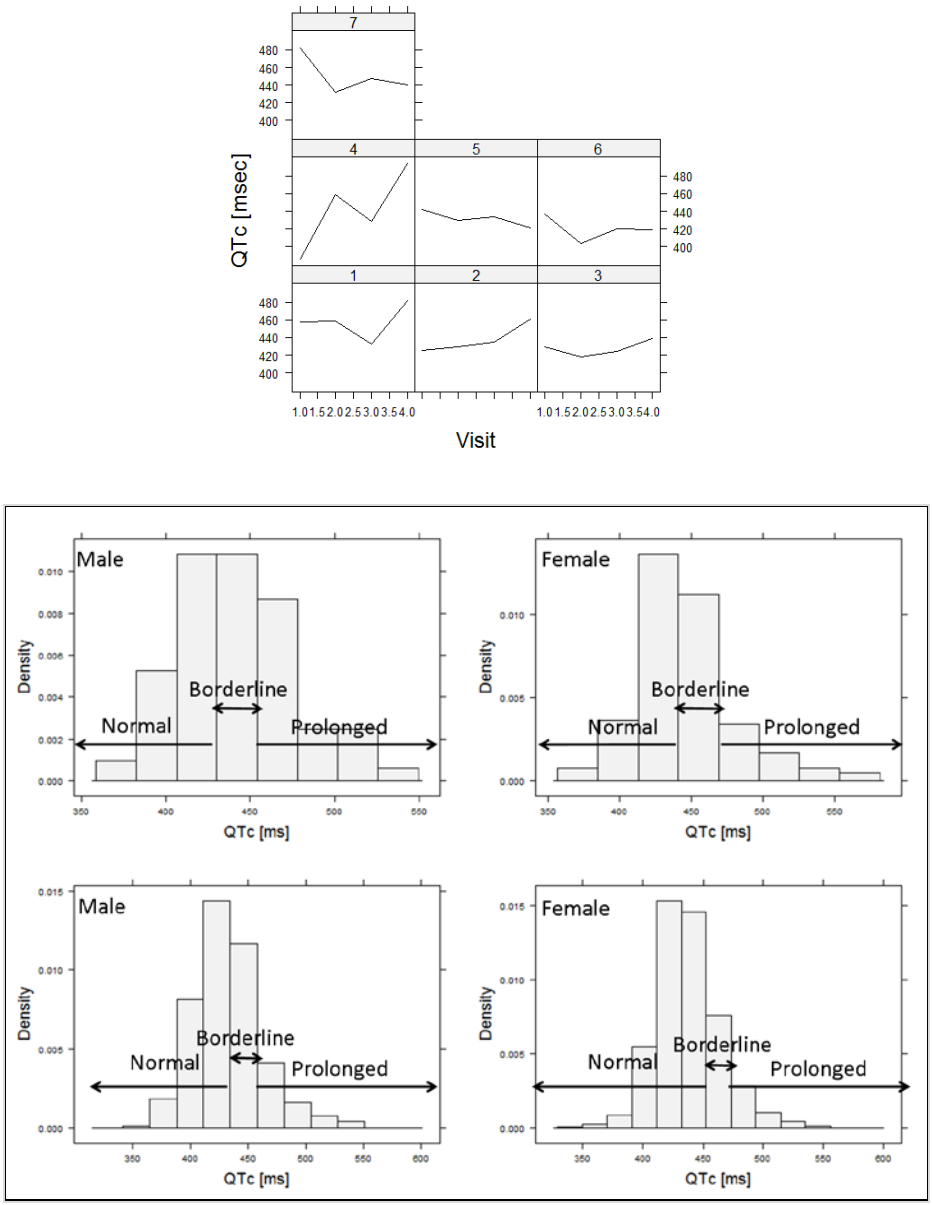


Figure 1. Four consecutive ECG measurements of seven randomly selected subjects prior to SCD (top panel). The x-axis denotes the four measurements (visit) between 1990-2004. Density distribution of last ECG measurements before SCD (middle panels). Density distribution of last ECG measurements of non SCD cases (bottom panels). The definition of "normal", "borderline", and "prolonged" QTc values are taken from the European regulatory guidelines. [19]

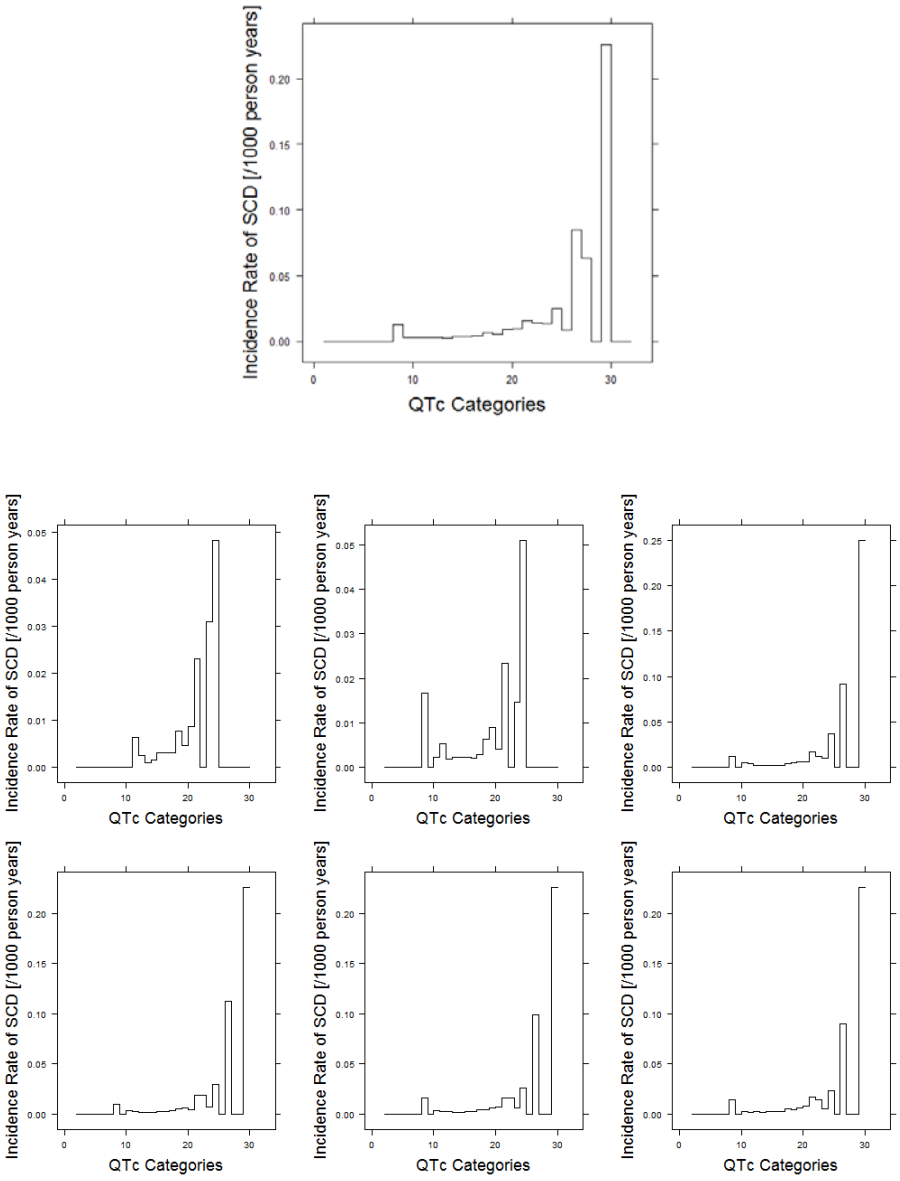


Figure 2. Top panel - Incidence rates of sudden cardiac death by increment bins of 10 msec QTc-intervals ($p < 0.00000235$). Bottom panels - Sensitivity analysis of incidence rates of sudden cardiac death. Each sub-panel shows the time lapse between the last measurement and death. N denotes the number of subjects included in the analysis.

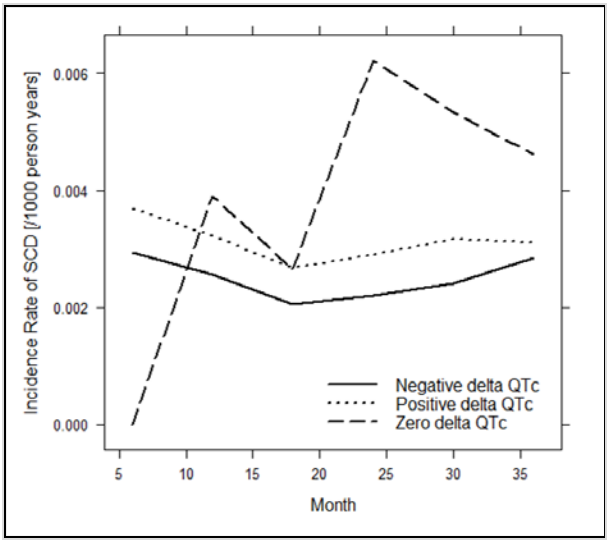


Figure 3. Incidence rates of sudden cardiac death by defining delta QTc as decreasing (negative delta), maintained (zero delta) and increasing (positive delta) (p = 0.195). The x-axis shows the time lapse between the last measurement and death.

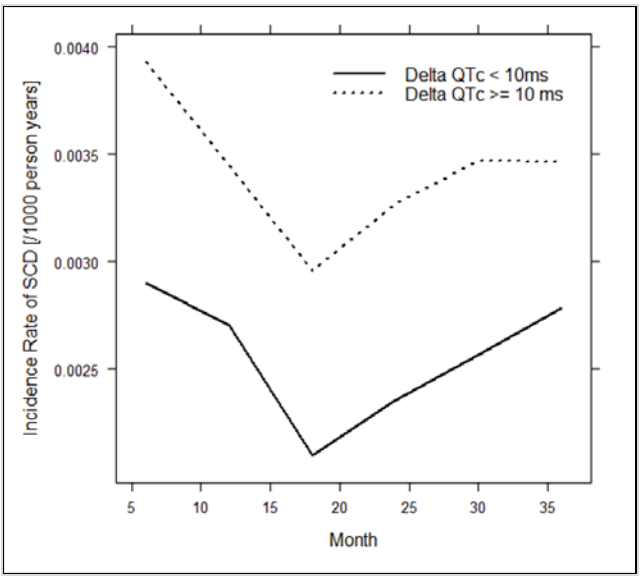


Figure 4. Incidence rates of sudden cardiac death by defining ΔQTc as < 10 msec and > 10 msec (p = 0.261). The x-axis shows the time lapse between the last measurement and death.

8.4 DISCUSSIONS

A prolonged QT-interval has been considered an attractive non-invasive risk factor for SCD. In this study, we showed that the incidence rate of SCD is better correlated with high absolute baseline QTc-interval values. A prolongation by ≥ 10msec did not significantly affect the incidence rate of SCD given there were only 182 cases with two consecutive QTc measurements prior to death.

Even though no consistent evidence has been demonstrated for increased risk of total or cardiovascular mortality or sudden death, the results from this study do agree with previous epidemiological studies. Recently, Straus *et al.*, found that prolonged QTc-intervals are associated with threefold increased risk of sudden cardiac death after adjustment for age, sex, body mass index, hypertension, cholesterol/HDL ratio, diabetes mellitus, myocardial infarction and heart failure in the same population used in this present study.[17] In another study, a QTc-interval > 440 msec was found to double the risk of sudden cardiac death.[15] Analogous conclusions have been drawn by the Amsterdam study and by Dekker *et al.*[24] Furthermore, inferences can also be made from data showing that the relative risk for arrhythmic death is higher when resting heart rate is greater or equal to 84 bpm[25], and that TdP is rarely associated with QTc-intervals less than 500 msec.[26]

On the other hand, our findings contrast with the Framingham study[11], which did not demonstrate an association of baseline QTc with sudden death. However, the population was considerably younger at time of enrolment (ages 30-65 years) and thus, the majority of baseline ECGs were recorded in patients probably before changes in risk become significant or cardiac disease is still subclinical. Other attempts to characterise the determinants of SCD have been less successful due to the limited number of SCD cases[16] or because of issues with the sensitivity and specificity of the analysis[27].

Despite the lack of evidence corroborating the link between QT/QTc-interval prolongation and TdP, QT-interval has been used as a surrogate marker for cardiovascular safety risks. When corrected for heart rate, the QT-interval read from an electrocardiogram is a direct marker of ventricular action potential duration and represent the time taken for the heart to return to its resting state following the preceding contraction that pumps blood around

the body. Albeit imperfect, it is at present the best available surrogate maker for the risk of TdP, which can potentially lead to SCD. From the sensitivity analysis performed in our study, SCD incidence rates were found to be independent from the time elapsed since last ECG measured and death. This implies that using the absolute QTc values to determine the risk of SCD may be relatively more reliable. [28,29]

Most importantly, our findings appear to contradict 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. From our analysis, the association between a delta 10 msec change in QTc and SCD is not statistically significant, nor reliable. Thus, two important concepts will need to be addressed in future revisions of the ICH E14 guideline: the threshold of 10 msec and the usage of delta QTc values.

There are also challenges in interpreting the significance of mean changes (ΔQT) from baseline without understanding inter-individual differences in drug exposure.[30] For example, the mean increases in peak QTc-interval were 9 and 22 msec following single oral doses of, respectively, 10 mg and 50 mg thioridazine[31], whilst 15 msec and 17 msec increases relative to placebo were found with 400 mg and 800 mg of single oral doses of moxifloxacin. [32] In this context, the assumption of a constant increase in the risk of pro-arrhythmia does not seem to be substantiated by clinical data. In fact, pharmacokinetic-pharmacodynamic analysis reveals clear differences in the slope of the concentration-effect relationships of compounds showing QTc-prolonging effects. [33]

Thus far, the risk of TdP as well as SCD has all pointed to an association with absolute QTc measurements. Based on our results, it appears that drug-induced changes in QTc may not increase the cardiovascular risks unless the measured QTc value falls within a range of values with a higher risk. [34]

It should be noted that our estimates were not adjusted for other cardiovascular risk factors, such as age, gender, smoking status, body mass index, diabetes and cardiac-related comorbidities, and that we used the most recent regulatory guidelines to classify QTc prolongation.[35] We have assumed random distribution in the measurement error and negligible confounding effect of drug-induced T-wave changes, which may complicate

the measurement of the QT-interval. Yet, it should be emphasised that the absolute risk of sudden cardiac death does show an increase with age. [36,37]

In summary, our findings corroborate an association between heart-rate corrected (QTc)-interval and the risk of sudden cardiac death (SCD). Given the growing relevance of QTc prolongation in drug development, our results clearly raise questions about the rationale for the evaluation of the pro-arrhythmic potential of novel compounds. Whilst it may 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 seems evident that clinically effective risk management policies should consider baseline QT/QTc-interval as a strong prognostic factor for pro-arrhythmias. We strongly recommend revisiting the current guidelines for the evaluation of cardiovascular safety. The assumption of a comparable risk of TdP across the population for an increase in QTc-interval of $\geq$10 msec cannot be sustained.

ACKNOWLEDGEMENTS

We would like to thank Dr. Jan Kors from the Department of Medical Informatics, Dr. Bruno Stricker, Dr. Charlotte van Noord and Dr. Jacqueline Witteman from the Department of Epidemiology at the Erasmus University Medical Centre for their invaluable contribution to the data used for the work presented here.

REFERENCES

1. *Al-Khatib SM, LaPointe NM, Kramer JM, Califf RM. What clinicians should know about the QT interval. JAMA. Apr 23-30 2003;289(16):2120-2127.*
2. *Wong SH, Mulvihill NT, Norton M. Assessing the risk of sudden cardiac death. Heart. Dec 2001;86(6):624-625.*
3. *Haverkamp W, Breithardt G, Camm AJ, et al. The potential for QT prolongation and proarrhythmia by non-antiarrhythmic drugs: clinical and regulatory implications. Report on a policy conference of the European Society of Cardiology. Eur Heart J. Aug 2000;21(15):1216-1231.*
4. *Myerburg RJ CA. Cardiac arrest and sudden cardiac death. . In: E B, ed. Heart disease: a textbook of cardiovascular medicine. New York: WB Saunders Publishing Co; 1997:742-779.*
5. *Priori SG, Aliot E, Blomstrom-Lundqvist C, et al. Task Force on Sudden Cardiac Death of the European Society of Cardiology. Eur Heart J. Aug 2001;22(16):1374-1450.*
6. *Killeen MJ. Drug-induced arrhythmias and sudden cardiac death: implications for the pharmaceutical industry. Drug Discov Today. Jun 2009;14(11-12):589-597.*
7. *Roden DM. Drug-induced prolongation of the QT interval. N Engl J Med. Mar 4 2004;350(10):1013-1022.*
8. *de Bruyne MC, Hoes AW, Kors JA, Hofman A, van Bemmel JH, Grobbee DE. Prolonged QT interval predicts cardiac and all-cause mortality in the elderly. The Rotterdam Study. Eur Heart J. Feb 1999;20(4):278-284.*
9. *Dekker JM, Schouten EG, Klootwijk P, Pool J, Kromhout D. Association between QT interval and coronary heart disease in middle-aged and elderly men. The Zutphen Study. Circulation. Aug 1994;90(2):779-785.*
10. *Elming H, Holm E, Jun L, et al. The prognostic value of the QT interval and QT interval dispersion in all-cause and cardiac mortality and morbidity in a population of Danish citizens. Eur Heart J. Sep 1998;19(9):1391-1400.*
11. *Goldberg RJ, Bengtson J, Chen ZY, Anderson KM, Locati E, Levy D. Duration of the QT interval and total and cardiovascular mortality in healthy persons (The Framingham Heart Study experience). Am J Cardiol. Jan 1 1991;67(1):55-58.*

12. *Montanez A, Ruskin JN, et al.. Prolonged QTc interval and risks of total and cardiovascular mortality and sudden death in the general population: a review and qualitative overview of the prospective cohort studies. Arch Intern Med. May 10 2004;164(9):943-948.*

13. *Okin PM, Devereux RB, et al. Assessment of QT interval and QT dispersion for prediction of all-cause and cardiovascular mortality in American Indians: The Strong Heart Study. Circulation. Jan 4-11 2000;101(1):61-66.*

14. *Robbins J, Nelson JC, Rautaharju PM, Gottdiener JS. The association between the length of the QT interval and mortality in the Cardiovascular Health Study. Am J Med. Dec 15 2003;115(9):689-694.*

15. *Algra A, Tijssen JG, Roelandt JR, Pool J, Lubsen J. QTc prolongation measured by standard 12-lead electrocardiography is an independent risk factor for sudden death due to cardiac arrest. Circulation. Jun 1991;83(6):1888-1894.*

16. *Karjalainen J, Reunanen A, Ristola P, Viitasalo M. QT interval as a cardiac risk factor in a middle aged population. Heart. Jun 1997;77(6):543-548.*

17. *Straus SM, Kors JA, De Bruin ML, et al. Prolonged QTc interval and risk of sudden cardiac death in a population of older adults. J Am Coll Cardiol. Jan 17 2006;47(2):362-367.*

18. *Fermini B, Fossa AA. The impact of drug-induced QT interval prolongation on drug discovery and development. Nat Rev Drug Discov. Jun 2003;2(6):439-447.*

19. *The assesment of the potential for QT interval prolongation by non-cardiovascular medicinal products. London: Human Medicines Evaluation Unit; 1997.*

20. *Assessment of the QT Prolongation Potential of Non-Antiarrhythmic Drugs. In: Heath Canada TPD, ed2001.*

21. *Harmonisation ICo. The clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs2005.*

22. *Hofman A, Grobbee DE, de Jong PT, van den Ouweland FA. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. Eur J Epidemiol. Jul 1991;7(4):403-422.*

23. *Hofman A, Breteler MM, van Duijn CM, et al. The Rotterdam Study: objectives and design update. Eur J Epidemiol. 2007;22(11):819-829.*

24. *Dekker JM, Crow RS, Hannan PJ, Schouten EG, Folsom AR. Heart rate-corrected QT interval prolongation predicts risk of coronary heart disease in black and white middle-aged men and women: the ARIC study. J Am Coll Cardiol. Feb 18 2004;43(4):565-571.*

25. *Camm AJ, Pakrashi, T., Savelieva, Irina. Sudden Cardiac Death: risk factors, treatment, and prevention. Dialogues in Cardiovascular Medicine. 2006;11(3):175-201.*

26. *Bednar MM, Harrigan EP, Anziano RJ, Camm AJ, Ruskin JN. The QT interval. Prog Cardiovasc Dis. Mar-Apr 2001;43(5 Suppl 1):1-45.*

27. *CIOMS. Current Cahllenges in Pharmacovigilance: Pragmatic Approaches. In: V. RoCWG, ed. Geneva2001.*

28. *Shah RR. Drugs, QT interval prolongation and ICH E14: the need to get it right. Drug Saf. 2005;28(2):115-125.*

29. *Gaita F, Giustetto C, Bianchi F, et al. Short QT Syndrome: a familial cause of sudden death. Circulation. Aug 26 2003;108(8):965-970.*

30. *Shah RR. Drug-induced prolongation of the QT interval: regulatory dilemmas and implications for approval and labelling of a new chemical entity. Fundam Clin Pharmacol. Apr 2002;16(2):147-156.*

31. *Hartigan-Go K, Bateman DN, Nyberg G, Martensson E, Thomas SH. Concentration-related pharmacodynamic effects of thioridazine and its metabolites in humans. Clin Pharmacol Ther. Nov 1996;60(5):543-553.*

32. *Demolis JL, Kubitza D, Tenneze L, Funck-Brentano C. Effect of a single oral dose of moxifloxacin (400 mg and 800 mg) on ventricular repolarization in healthy subjects. Clin Pharmacol Ther. Dec 2000;68(6):658-666.*

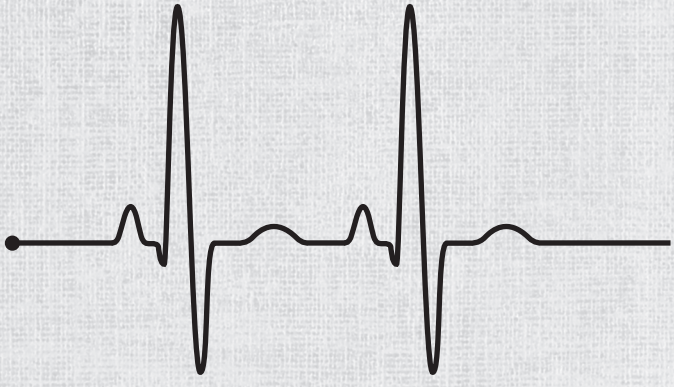
33. *Chain AS, Krudys KM, Danhof M, Della Pasqua O. Assessing the Probability of Drug-Induced QTc-Interval Prolongation During Clinical Drug Development. Clin Pharmacol Ther. Dec;90(6):867-875.*

34. *Shah RR. Drugs, QTc interval prolongation and final ICH E14 guideline : an important milestone with challenges ahead. Drug Saf. 2005;28(11):1009-1028.*

35. *Kardys I, Kors JA, van der Meer IM, Hofman A, van der Kuip DA, Witteman JC. Spatial QRS-T angle predicts cardiac death in a general population. Eur Heart J. Jul 2003;24(14):1357-1364.*

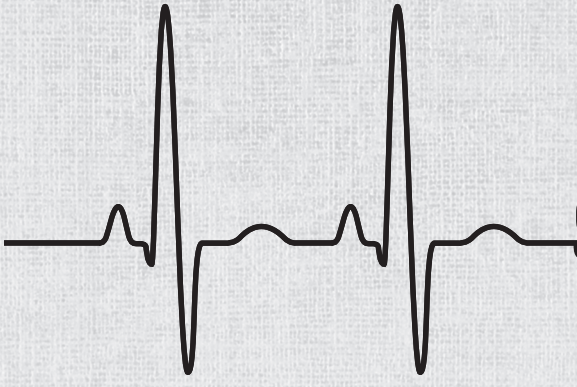
36. *Krahn AD, Connolly SJ, Roberts RS, Gent M. Diminishing proportional risk of sudden death with advancing age: implications for prevention of sudden death. Am Heart J. May 2004;147(5):837-840.*

37. *Zipes DP, Wellens HJ. Sudden cardiac death. Circulation. Nov 24 1998;98(21):2334-2351.*



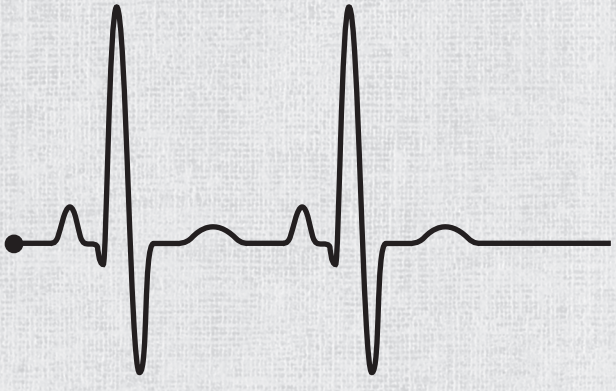
"Any intelligent fool can make things bigger and more complex... It takes a touch of genius - and a lot of courage to move in the opposite direction."

Albert Einstein



SECTION IV

CLINICAL TRIAL SIMULATION IN THE EVALUATION OF CARDIOVASCULAR SAFETY

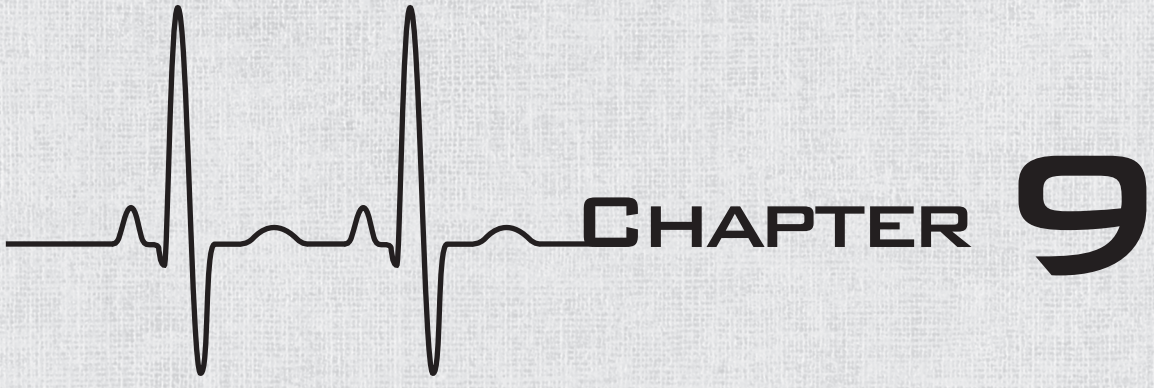


Model-based Simulation Of QT And RR Intervals For The Evaluation Of QTc-Interval Prolongation

ASY Chain (1), F Bellanti (1), M Danhof (1),
MCJM Sturkenboom (2, 3), OE Della Pasqua (1,4)

(1) Leiden/Amsterdam Centre for Drug Research,
Division of Pharmacology, Leiden University, Leiden, The Netherlands
(2) Department of Medical Informatics and Epidemiology,
Erasmus Medical Centre, The Netherlands
(3) Department of Epidemiology, Erasmus Medical Centre,
The Netherlands
(4) Clinical Pharmacology and Discovery Medicine,
GlaxoSmithKline, Stockley Park, Uxbridge, United Kingdom

British Journal of Clinical Pharmacolgy - in press



Chapter 9

ABSTRACT

Aim: Correction for changes in heart rate is a fundamental step to the evaluation of QTc-interval prolongation. Yet, clinical trial simulations for thorough-QT (TQT) studies often rely on approximations to evaluate design factors such as group size. The aim of the investigation was to develop a model-based approach to describe the correlation between QT- and RR-intervals in healthy volunteers.

Methods: A large pool of healthy volunteer ECG data has been used for model building and validation. The data was split into two subsets to allow for the external validation of the final model where analysis was performed using a nonlinear mixed effects approach. Model validation was carried out internally and externally.

Results: A power function gave the best model performance. Age and gender were the only available covariates; gender was found to be significant both on slope [181 (sd. 13.5) for females and 166 (sd. 12.8) for males] and exponent [0.85 (sd. 0.13) for females; 0.74 (sd. 0.13) for males]. Inter-occasion variability on slope and exponent was also identified as a significant random effect. Distributions of simulated and real QT values were comparable. Parameter estimates were subsequently used as part of a TQT study simulation.

Conclusions: The final model allows a reliable and realistic simulation of QT-interval profiles starting from a physiological set of RR values. In the context of clinical trial simulations, the availability of such a model represents a concrete improvement in the evaluation of drug-induced QTc-interval prolongation.

9.1 INTRODUCTION

The QT-interval in the electrocardiogram (ECG) represents the total duration of depolarisation and repolarisation of the cardiac cycle. Drug-induced prolongation of ventricular repolarisation, also known as QT prolongation, is considered the best surrogate marker of risk for ventricular arrhythmia [1]. The assessment of a drug's effect on the QT-interval has therefore become mandatory during drug development [2, 3]. However, heart rate 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 [4]. Correction for heart rate is necessary to provide a single intrinsic physiological value that can be compared between subjects and within the same subject under different conditions. 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.

In the past few years, a considerable number of mathematical functions has been described that supposedly best characterise the heart rate-QT relationship [5-12]. Whilst the focus of those investigations has been on the accuracy and precision of the parameters describing the so-called physiological relationship, little attention has been paid to the use of such functions in clinical trial simulations to support the design of clinical protocols aimed at the evaluation of drug-induced effects on QT/QTc-interval, and in particular of the thorough-QT (TQT) study [2]. The ability to decide if a new chemical entity (NCE) should be discontinued early in development depends on the implementation of a well-designed protocol, in which statistical power calculations take into account the non-random nature of variability in QT-interval. Similar considerations also apply to study design when concentration-effect relationships are used as basis for the assessment of drug-induced effects.

Recently, we have proposed the use of a Bayesian hierarchical model for the evaluation of QTc-interval prolongation using phase I data in lieu of a specific TQT study [13]. Effective implementation of this approach requires the use of clinical trial simulations (CTS) to ensure accurate assessment of drug-induced effect size and to determine if a TQT study is actually necessary. Simulations are also essential to enable the characterisation of type I and II error rates. Traditionally, however, re-sampling techniques from existing

data or empirical correlations have been used to evaluate drug effect and design factors such as group size.

The availability of a method to generate realistic RR profiles according to typical protocol sampling schemes, which can be used as input into a model describing the QT-RR relationship, would be very useful. In the present study, we aim to generate 24-hour RR profiles and develop a nonlinear mixed effects model to characterise QT-RR relationship in healthy subjects. To meet the aforementioned objectives, challenges must be overcome from both a physiological and a statistical perspective.

Non-linear regression methods have been applied, which rely on a fundamental assumption that the predictor variable (RR-interval) is measured without error [14]. However, the fact is that both QT- and RR-intervals measured in electrocardiogram (ECG) are subject to not only measurement error, but also fluctuation that is caused by physiological and biological factors [15]. In addition, statistically, it has been shown that the relationship between the RR-interval (=1/heart rate) and the QT-interval is highly variable between individuals but stable over time within each individual, but none of investigations has looked closely at the variability associated with the sampling scheme [16]. In contrast to previous research, in which Bayesian hierarchical model was used to account for or exclude measurement error in the evaluation of QTc-interval prolongation [13], we have chosen to simulate data in which correlations and variance structure reflect the noise due to measurement and sampling procedures.

9.2 METHODS

Data

The data for this study was derived from a large pool of healthy volunteer ECG data, totaling 776 subjects (339 males and 437 females), with mean age of 31.1 years (sd. 9.4). Variables such as QT, QTc, RR, VR, QRS, PR, age, weight and gender were available. From seven different studies, the 24-hour profiles of RR- and QT-intervals measured at discrete time points according to one of the seven protocols, were used in this analysis. The mean QT value and the mean RR value from the combined dataset were 390.5 msec (sd. 27.8) and 0.96 msec (sd. 0.15) respectively.

Generating RR-iIntervals

Using the starting RR-interval values (RR_0) from each of the 776 profiles, a vector of RR-intervals was created to sample from for a new hypothetical population at time 0. Then the median RR_0 was determined. Using the median value, two distributions of RR differences were defined, one each for above and below the median. We chose to split the values at the median because it was found that the two distributions statistically differ from each other whereas the distributions when the data is separated by bins of different sizes or quartiles were not significantly different. The definition of differences in RR values was defined according to the following formula:

$$\Delta RR1 = M1 - M2; \Delta RR2 = M1 - M3; \ldots, \text{where}$$

M1 = first RR measurement, M2 = second RR measurement, etc.

After the RR value for time 0 for each of the profile in the hypothetical population was sampled, the RR differences for the remaining profile were sampled from one of the two distributions described earlier, depending on the starting value. Finally, a full 24-hour RR profile was back-calculated for each subject based on the differences sampled, according to a typical clinical trial sampling scheme as shown in Table 1. Since the new RR profiles generated were sampled non-parametrically, physiological boundaries were used to prevent unrealistic values.

	Baseline				Dose						Post-dose						
Time	-1	-0.75	-0.5	-0.25	0	0.5	1	1.5	2	2.5	3	4	6	8	12	18	24
PK	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x	x
PD	x		x	x		x	x	x	x			x		x	x		x

Table 1. Typical clinical trial sampling scheme

Generating QT-intervals

QT-intervals for each RR-interval generated according to the aforementioned steps were simulated according to the following relationship,

$$\mathbf{QT = SLP * RR\text{^}EXP + INCPT},$$

between all the pairs of QT and RR values in the healthy subject data, where SLP is the slope parameter of the relationship, EXP is the exponential term quadratic function and INCPT is the parameter of the y-intercept. The data was split into two subsets to allow for model building and external validation of the final model. Model building was performed using nonlinear mixed effects approach as implemented in NONMEM VI (ICON Development Solutions**,** Ellicott City, MD, USA) [17].

The criteria of model selection were based on changes in objective function (OFV; equal to -2 log likelihood) and goodness of-fit plots (GOF). Gender, age and weight were tested as covariates on the parameters SLP, EXP and INCPT but only gender was found to have an impact on SLP and EXP. In addition, visual predictive checks (VPC) and normalised prediction distribution error (NPDE) plots were used to judge model performance with regard to the predictions for a different subset of the data.

Goodness-of-fit plots, including observed (OBS) *versus* individual prediction (IPRED), OBS *versus* population prediction (PRED), conditional weighted residuals (CWRES) [18] *versus* time and CWRES *versus* OBS were used for diagnostic purposes. The NPDE method was used to determine whether 1000 simulated QT values were interchangeable with the original data using graphical diagnostics [19]. NPDE was implemented with an add-on software package in R2.8 (The R project - http://www.r-project.org/) [20].

9.3 RESULTS

RR-intervals

In general, all simulated values are visually compared to the real data by overlapping the distributions to determine if the simulated values resemble the original data. The resulting distributions of RR differences are shown in Figures 1. After the hypothetical profiles of an entire population were simulated, the distributions of actual RR values are displayed in Figures 2.

Actual values, population and individual profiles, are also inspected visually to determine if the variability of RR profiles are consistent over time between real and simulated data. Population comparisons are shown in Figure 3 and Figure 4 are individual profiles.

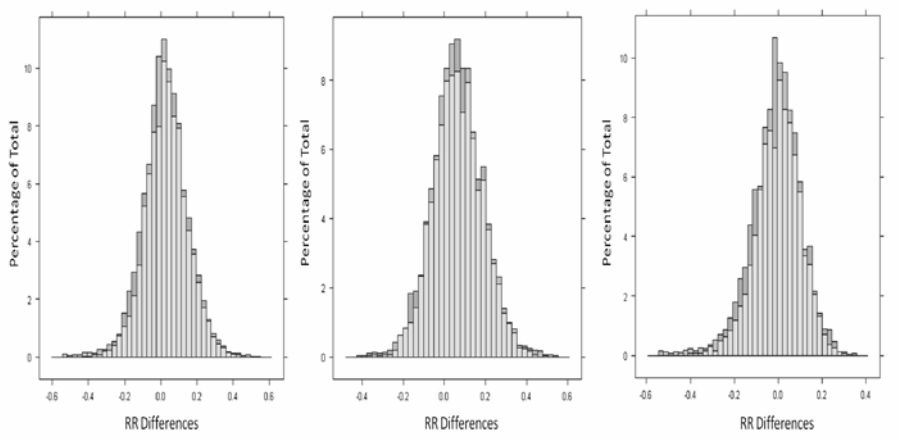


Figure 1. Distribution of RR differences of the entire population (left), above the median (middle) and below the median (right). The lighter colour in the histogram represents the portion of overlapped data between the observed and the simulated values.

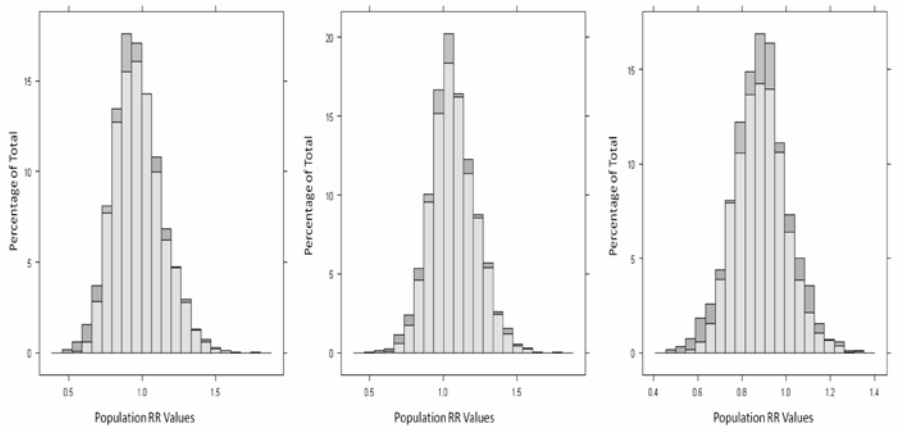


Figure 2. Distribution of RR values of the entire population (left), above the median (middle), and below the median (right). The lighter colour in the histogram represents the portion of overlapped data between the observed and the simulated values.

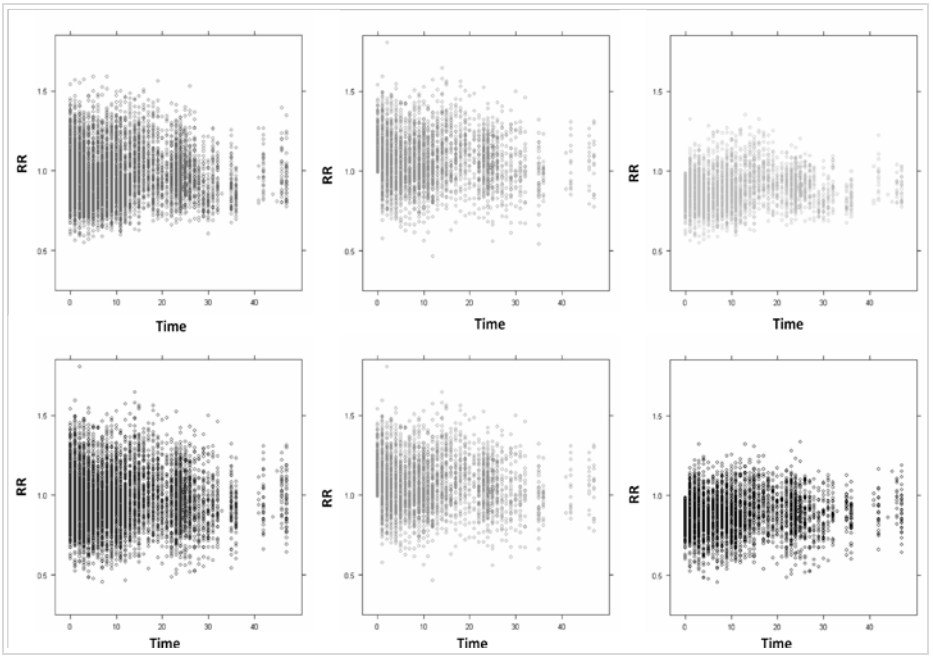


Figure 3. Distribution of RR values over time. Top panels: all real data (left), above the median (middle) and below the median (right). Bottom panels: all simulated data (left), above the median (middle) and below the median (right).

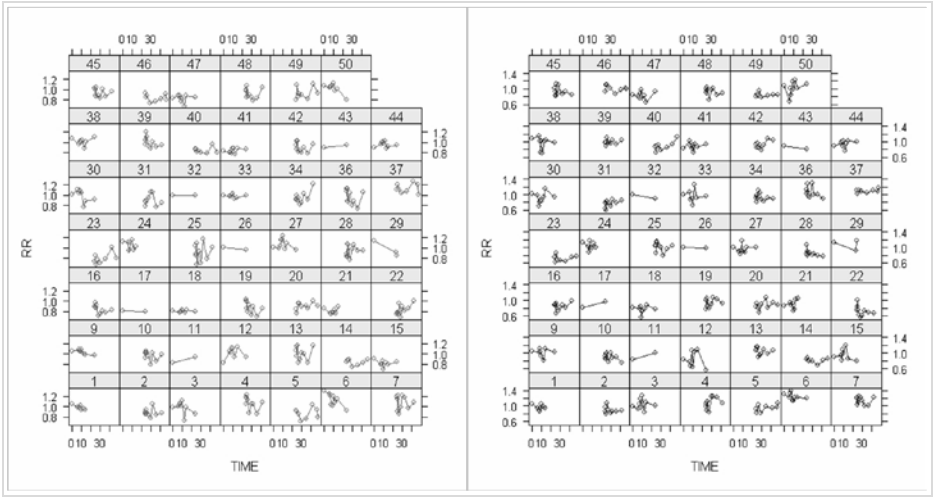


Figure 4. Individual profiles of RR-intervals with observed data (left) and simulated data (right).

QT-intervals

Among the different functions used in the evaluation of the QT-RR correlation [21], a power function yielded the best model performance. The final model parameters are shown in Table 2. Gender was found to be a significant covariate both on slope and exponent. Inter-occasion variability on slope (SLP) and exponent (EXP) was also identified as a significant random effect**.** Model diagnostics of the final model are shown in Figure 5, which provided the confidence that the model is able to well describe the data. Figure 6 is the model evaluations of the external validations using a subset of the original data.

	SLP			**EXP**			**INCPT**	**IOV**	**Residual**
	female	male	IIV	female	male	IIV			
Model estimate	181	166	0.0059	0.85	0.74	0.046	221	0.012	0.0004
Bootstrap mean	180.2	165.3	0.0062	0.87	0.76	0.0459	222.3	0.013	0.0004
CV %	12.3	13.7	29	13.7	13.1	17.4	10.1	38	5

Table 2. Final model estimates for the QT-RR correlation.

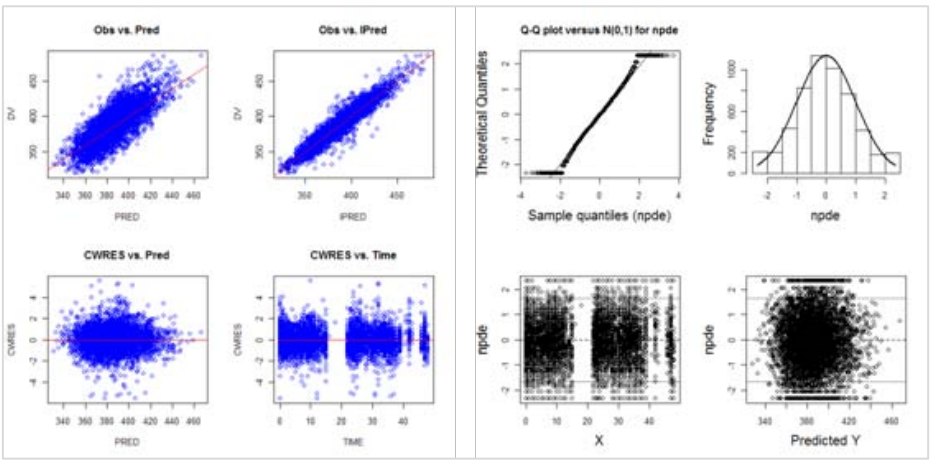


Figure 5. Right panel: goodness-of-fit plots and left panel: NPDE, for final QT-RR model evaluation.

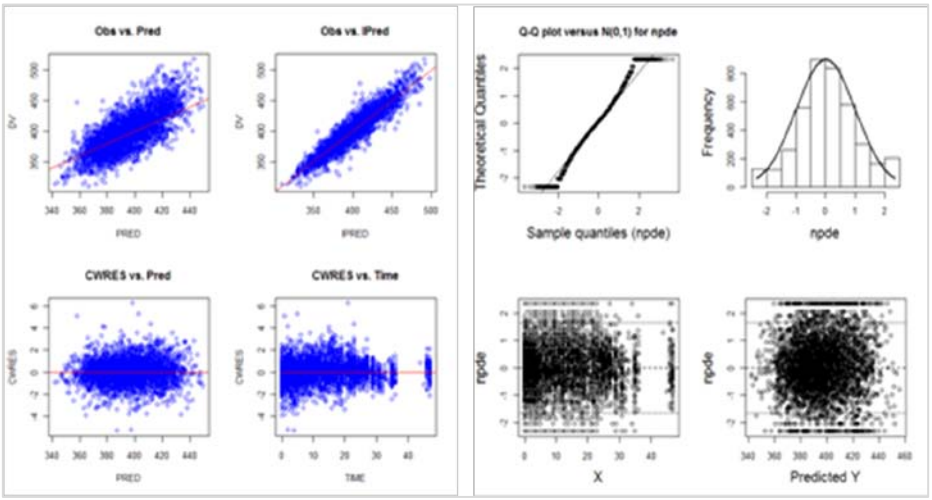


Figure 6. Right panel: goodness-of-fit plots and left panel: NPDE, for external validation of final QT-RR mode using an external dataset.

QT values were simulated for the RR-intervals. Figure 7 shows the overlapping distributions between real and simulated QT data as well as the comparison between the individual profiles.

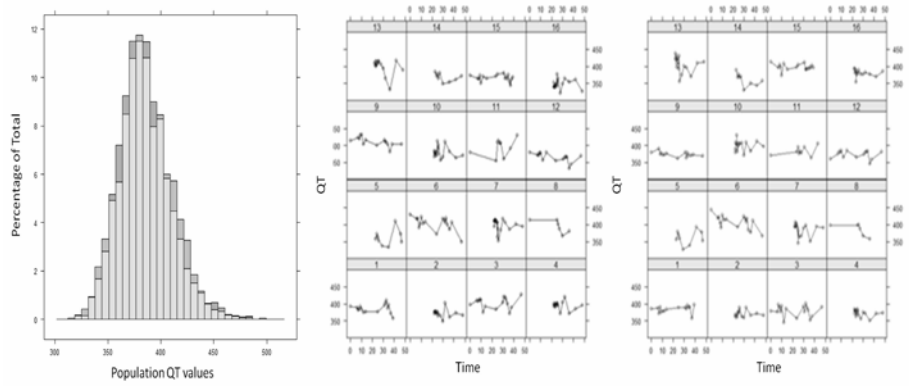


Figure 7. Overlapping distributions of real and simulated QT values (left), individual profiles of observed QT data (middle) and simulated profiles (right) of QT-intervals. The lighter colour in the histogram represents the portion of overlapped data between the observed and the simulated values.

9.4 DISCUSSIONS

Realistic QT-intervals were obtained, from a novel modelling approach describing the QT-RR relationship. Specifically, the developed model allowed for accurate simulation of QT-interval profiles starting form a physiological set of RR values. In the context of clinical trial simulations, the availability of such a model offers considerable advantages as compared to re-sampling methodologies. The use of a model-base method allows one to generate an infinite number of realistic individual QT profiles, which can be used subsequently in conjunction with a PKPD model to assess the liability of a compound, including QTc-interval prolongation. This approach also enables the prediction of potential drug effects on QTc-interval for patients who do not meet inclusion/exclusion criteria.

Various models have been proposed to simulate beat-to-beat RR-interval [22, 23]; however, they are not necessarily useful when RR measurements are needed for clinical trial simulations (CTS). To apply RR-intervals in CTS, measurements should be taken at discrete intervals according to the study protocol, yielding between-measurement variability which is not described by beat-to-beat variation. Our method represents a valuable tool for the optimisation of drug development processes, especially in studies such as TQT.

Recently, Chen et al proposed a Bayesian hierarchical measurement error modelling approach to assess QT-interval correction and prolongation [14]. The novel method not only takes into account the measurement errors in RR-intervals, but it also provides a unified approach to QT-interval correction and prolongation evaluation in place of other methods to correcting QT-interval for heart rate [24, 25]. This method would be useful when assessing drug-induced QTc prolongation of NCE, however, it would not be useful towards other clinical trial simulations that do not aim to determine pro-arrhythmic propensity of compounds.

Conversely, our newly proposed approach to non-parametric sampling of RR-intervals does not take into account any possible measurement errors stemming from the ECG recordings. Furthermore, this approach assumes all RR-interval values fall within the same distribution which is not always the case when subjects are in different treatment arms or dose groups. Thus the

pharmacological effect on RR-intervals of the compound under investigation must be well understood prior to using the non-parametric methodology.

One advantage to using the approach presented here is that a fixed number of subjects with a number of observations can provide the basis for a large scale simulation. Thus as long as the sampling intervals remain the same between observed data and sampled data, there would not be any bias in the new dataset. By extension, this method also offers advantages over the traditional bootstrapping technique. Chief of which is the fact that `new` subjects can be created with entirely different observations as opposed to a re-sampling approach where new sets of data are taken from the observed information using sampling with replacement so that it is not identical to the original set.

Potential applications for this novel method to simulating discrete RR measurements include any investigations involving PR-intervals, ST-intervals, QRS-intervals where changes in RR-intervals matter [26]. Other physiological research in which RR is the variable of interest, such as anti-inflammatory pathways [27], respiratory functions [28, 29] and cardiovascular safety [30] can also benefit from this approach. Thus the usefulness of the technique discussed in the present study goes far beyond thorough-QT study and other clinical trial simulations.

REFERENCES

1. *Roden DM. Drug-induced prolongation of the QT interval. N Engl J Med 2004; 350: 1013-22.*

2. *Guidance for Industry: ICH E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs. U.S. Department of Health and Human Services. Food and Drug Administration. (2005.)*

3. *Markert M, Shen R, Trautmann T, Guth B. Heart rate correction models to detect QT interval prolongation in novel pharmaceutical development. J Pharmacol Toxicol Methods 2011; 64: 25-41.*

4. *Davey P. How to correct the QT interval for the effects of heart rate in clinical studies. J Pharmacol Toxicol Methods 2002; 48: 3-9.*

5. *Batchvarov V, Malik M. Individual patterns of QT/RR relationship. Card Electrophysiol Rev 2002; 6: 282-8.*

6. *Batchvarov VN, Ghuran A, Smetana P, Hnatkova K, Harries M, Dilaveris P, Camm AJ, Malik M. QT-RR relationship in healthy subjects exhibits substantial intersubject variability and high intrasubject stability. Am J Physiol Heart Circ Physiol 2002; 282: H2356-63.*

7. *Bazett JC. An analysis of time relations of electrocardiograms. Heart 1920; 7: 353-67.*

8. *Desai M, Li L, Desta Z, Malik M, Flockhart D. Variability of heart rate correction methods for the QT interval. Br J Clin Pharmacol 2003; 55: 511-7.*

9. *Fridericia LS. Die systolendauer im elekrokardiogramm bei normalen menschen und bei herzkranken. Acta Med Scand 1920; 53.*

10. *Malik M. The imprecision in heart rate correction may lead to artificial observations of drug induced QT interval changes. Pacing Clin Electrophysiol 2002; 25: 209-16.*

11. *Malik M, Hnatkova K, Batchvarov V. Differences between study-specific and subject-specific heart rate corrections of the QT interval in investigations of drug induced QTc prolongation. Pacing Clin Electrophysiol 2004; 27: 791-800.*

12. *Sagie A, Larson MG, Goldberg RJ, Bengtson JR, Levy D. An improved method for adjusting the QT interval for heart rate (the Framingham Heart Study). Am J Cardiol 1992; 70: 797-801.*

13. *Chain AS, Krudys KM, Danhof M, Della Pasqua O. Assessing the Probability of Drug-Induced QTc-Interval Prolongation During Clinical Drug Development. Clin Pharmacol Ther 2011; 90: 867-75.*

14. *Chen J, Zhao X. A Bayesian measurement error approach to QT interval correction and prolongation. J Biopharm Stat 2010; 20: 523-42.*

15. *Smetana P, Batchvarov V, Hnatkova K, Camm AJ, Malik M. Circadian rhythm of the corrected QT interval: impact of different heart rate correction models. Pacing Clin Electrophysiol 2003; 26: 383-6.*

16. Jacquemet V, Dube B, Knight R, Nadeau R, LeBlanc AR, Sturmer M, Becker G, Vinet A, Kus T. Evaluation of a subject-specific transfer-function-based nonlinear QT interval rate-correction method. Physiol Meas 2011; 32: 619-35.

17. Beal S, Sheiner, L. . NONMEM user's guide In: NONMEM project group, San Francisco, 1999.

18. Hooker AC, Staatz CE, Karlsson MO. Conditional weighted residuals (CWRES): a model diagnostic for the FOCE method. Pharm Res 2007; 24: 2187-97.

19. Comets E, Brendel K, Mentre F. Computing normalised prediction distribution errors to evaluate nonlinear mixed-effect models: the npde add-on package for R. Comput Methods Programs Biomed 2008; 90: 154-66.

20. R Development Core Team, R: a language and environment for statistical compting. In, edcomputing Rffs, Vienna, Austria, 2010.

21. Malik M, Farbom P, Batchvarov V, Hnatkova K, Camm AJ. Relation between QT and RR intervals is highly individual among healthy subjects: implications for heart rate correction of the QT interval. Heart 2002; 87: 220-8.

22. Bowers E.J. L, P., Drinnan, M.J., Smith, F.E., Allen, J., Murray, A. Simulation of cardiac interval time series. In: Computers in Cardiology: IEEE, 2002: 233-36.

23. Mandrekar SJ, Nagaraja HN, Berntson GG. Statistical modelling of the differences between successive R-R intervals. Stat Med 2005; 24: 437-51.

24. Ring A. Statistical models for heart rate correction of the QT interval. Stat Med 2010; 29: 786-96.

25. Wang D, Cheung YB, Arezina R, Taubel J, Camm AJ. A nonparametric approach to QT interval correction for heart rate. J Biopharm Stat 2010; 20: 508-22.

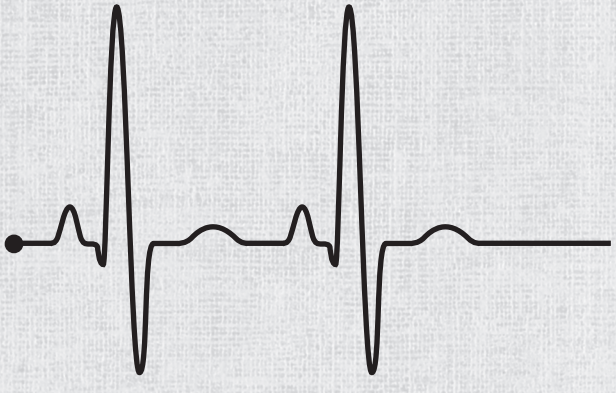
26. Malik M, Hnatkova K, Sisakova M, Schmidt G. Subject-specific heart rate dependency of electrocardiographic QT, PQ, and QRS intervals. J Electrocardiol 2008; 41: 491-7.

27. Sloan RP, McCreath H, Tracey KJ, Sidney S, Liu K, Seeman T. RR interval variability is inversely related to inflammatory markers: the CARDIA study. Mol Med 2007; 13: 178-84.

28. Brown TE, Beightol LA, Koh J, Eckberg DL. Important influence of respiration on human R-R interval power spectra is largely ignored. J Appl Physiol 1993; 75: 2310-7.

29. Milic M, Bao X, Rizos D, Liu F, Ziegler MG. Literature review and pilot studies of the effect of QT correction formulas on reported beta2-agonist-induced QTc prolongation. Clin Ther 2006; 28: 582-90.

30. van Noord C, Sturkenboom MC, Straus SM, Hofman A, Kors JA, Witteman JC, Stricker BH. Serum glucose and insulin are associated with QTc and RR intervals in nondiabetic elderly. Eur J Endocrinol 2010; 162: 241-8.

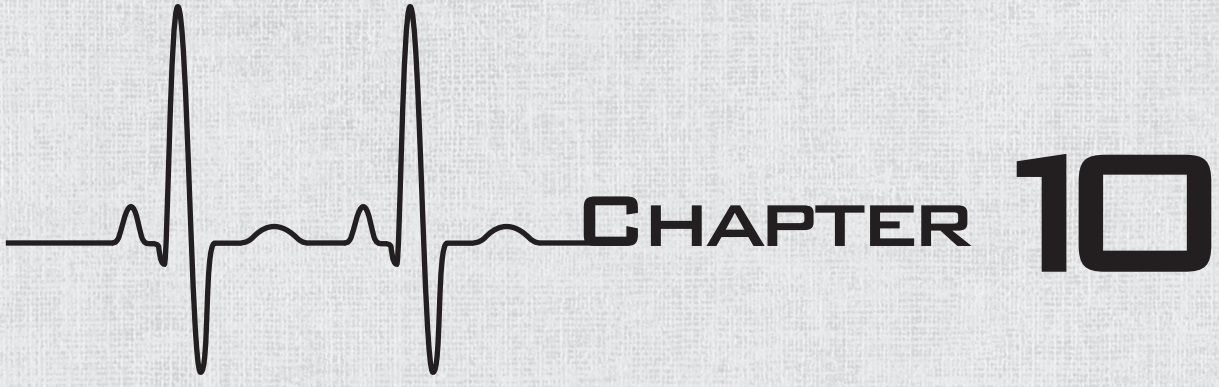


Model-based Analysis Of Thorough QT Studies

ASY Chain (1), F Bellanti (1), M Danhof (1),
MCJM Sturkenboom (2, 3), OE Della Pasqua (1,4)

(1) Leiden/Amsterdam Centre for Drug Research,
Division of Pharmacology, Leiden University, Leiden, The Netherlands
(2) Department of Medical Informatics and Epidemiology,
Erasmus Medical Centre, The Netherlands
(3) Department of Epidemiology, Erasmus Medical Centre,
The Netherlands
(4) Clinical Pharmacology and Discovery Medicine,
GlaxoSmithKline, Stockley Park, Uxbridge, United Kingdom

Submitted for publication



Chapter 10

ABSTRACT

The aim of the study was to investigate if the previously proposed Bayesian approach to characterise the probability of drug-induced QTc-interval prolongation can overcome the statistical issues associated with the recommended double-delta method for the analysis of thorough-QT studies. A completely simulated trial population was created based on non-parametric RR samplings and its corresponding QT-intervals using hierarchical modelling technique. Thirty-six different TQT studies were simulated with a hypothetically fixed drug-effect size, trial design and numbers of subjects, To compare the two different approaches, sensitivity and specificity were calculated and tabulated for each method to capture false negative and false positive rates. In general, false positive trials were only detected with the double-delta method and lower specificity was observed with smaller sample size. The results also showed the Bayesian approach is able to detect QTc-interval prolongation with higher accuracy and precision with different number of subjects and various trial designs.

10.1 INTRODUCTION

In 2005, the International Conference on Harmonisation (ICH) published a guidance, topic E14, for clinical evaluation of QT prolongation as a result of increasing concern for drug-induced proarrhythmic events [1]. Since QT-interval prolongation has been considered a surrogate marker for risk of developing the potentially fatal tachyarrhythmia, torsade de pointes (TdP), the guidance advocated the use of thorough-QT (TQT) studies to evaluate the propensity of QT/QTc-interval prolongation caused by a new chemical entity (NCE). TQT studies, as recommended, should be conducted early in clinical development to exclude compounds that produce a prolongation exceeding a threshold when comparing active to placebo treatment arms. A study is only deemed negative when the upper limits of all the 1-sided 95% confidence intervals (CI) are below 10 milliseconds at every time point [2]. Conversely, a study is interpreted as positive if just one CI includes a prolongation of 10 milliseconds or greater.

Presently these studies must be powered for the statistical detection of QTc-interval changes as small as 5 milliseconds. Since this translates into a noticeable drug development burden, strategies need to be identified allowing the size and thus the cost of thorough-QT/QTc studies to be minimised. Efforts have been made to investigate the influence of QT- and RR-interval data quality and the precision of heart rate correction on the sample sizes of thorough-QT/QTc studies [3]. Despite these concerns, research focus has been restricted to the investigation of correction factors and measurement techniques as determinants of precision and accuracy [3, 4]. Instead, more recently, it has become clear that the choice of the data analysis method plays an equally or even more important role in the precision and accuracy of results.

In recent research, the analysis of baseline values, study sensitivity with positive control arm and formulas to correct QT-intervals for heart rate, have been explored in detail as they have been determined to impact the precision and accuracy of results [5]. Baseline measurement, if not controlled and consistent across studies, could lead to significant misinterpretation of results as drug, placebo and active control are estimated relative to it [6]. In an analysis conducted by Sethuraman et al., it was determined that the choice of baseline approaches impacts the type I error and power for various study designs [7]. Furthermore, there is a high probability of concluding that

the TQT study is positive when the two groups differ considerably at baseline in a study with parallel design [8]. An equally important consideration is the use of a positive control. The purpose of including a positive control arm is to demonstrate the study's ability to detect a small effect on the QT-interval. In most studies, moxifloxacin has been selected as it supports the appropriateness of its use in TQT studies [9]. Finally, QT correction for heart rate has also been an important topic of discussion when interpreting TQT study results [10]. Since the Bazett's and Fridericia's formula from 1920, many new ways to correct QT for heart rate have been proposed such as a study-population-specific correction formula [11] and individual subject-specific correction formula [12].

From a statistical perspective, the suggested method of analysis in the ICH guidance is the double-delta method where the time-matched mean QTc-interval difference between active and placebo treatments, both adjusted for baseline, is taken [13]. Then an assessment of non-inferiority of the supra-therapeutic dose to placebo is performed by the intersection-union test [14]. Thus the statistical method used to calculate sample size on power depending on the covariance structure of the data needs to be chosen carefully [14-16]. Methods of ECG measurements, such as manual vs. automated [17], and statistical approaches used to analyse inter-and intra-reader variability for ECG measurements also requires attention as the traditional Bland-Altman method [18-21] may not be applicable when more than two readers are involved [4]. Ultimately, the careful considerations in using various statistical methods in different aspects of a TQT study are to ensure a well balanced specificity and sensitivity of results.

A comprehensive evaluation of type I and II error rates is highly desirable given the clinical and financial impact of inaccurate decisions regarding drug-induced QTc interval prolongation. Moreover, alternative methodology should be considered in addition to standard hypothesis testing. Frequentist methods aimed at detecting a threshold or cut-off value for the effect size are not clinically as relevant as methods which enable characterisation of how changes in drug exposure correlate with changes in QTc-interval, particularly if such effects are observed primarily at supra-therapeutic levels.

Currently, evidence shows that the double-delta method produces high false-positive rates [22]. These rates are also influenced by various study design factors, such as design type (i.e. crossover and parallel), sample size and number of baseline QT measurements taken [2]. Recently, we have shown the potential advantages of applying a Bayesian approach to characterise the probability of drug-induced QTc-interval prolongation and the underlying concentration-effect relationship [23]. Here we show the clinical and statistical implications of using a model-based approach to design and analyse drug-induced effects in a TQT study. Using simulations with a range of parameter estimates defining as negative (< 5 msec), borderline (~ 5msec) and positive (≥ 10 msec) drug effects, we assessed type I and II error rates for different population sizes and study designs using a hypothetical population of healthy subjects. We expect to substantiate previous findings which suggest that the use of the double-delta method is inappropriate for the analysis of TQT studies.

10.2 METHODS

SIMULATION OF EXPERIMENTAL VARIABLES IN A HYPOTHETICAL POPULATION

Assumptions

The steps taken to generate a dataset for a hypothetical population were not done without certain assumptions. For instance, we assumed that the random noise in the simulated data reflects the actual variability in measurements. In manually read ECG data, there exist between-observer variability and potentially other sources of bias due to the algorithm used to harmonise the differences, while machine-read ECG values also include measurement errors.

No demographic or other covariates were simulated for the hypothetical population because these were deemed as random factors and as such will not interfere with the drug-induced effect size estimation. Furthermore, we made the assumption that there was no metabolite formation which could lead to delayed onset of effects. Only the parent drug affected QT-intervals and not RR, thus only minor within-subject variability existed in the QT/RR relationship.

The maximum QTc-interval prolongation had been set to occur at the same time as the maximum concentration. Therefore if the sampling frequency had been reduced, it can be assumed that the true drug-induced effect may be missed by any analytical method.

Data generation

An overview to generate a dataset for a hypothetical population to compare the methods for analysing QTc-interval prolongation is shown in Figure 1.

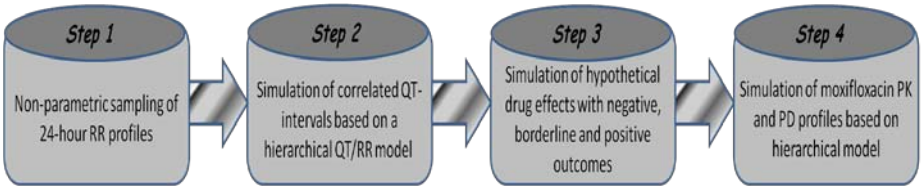


Figure 1. Flow diagram describing the steps taken to generate a full dataset for a hypothetical population.

Study population demographics

The data for this study was derived from a large pool of healthy volunteer ECG data, totaling 776 subjects (339 males and 437 females), with mean age of 31.1 years (sd. 9.4). Variables such as QT, QTc, RR, VR, QRS, PR, age, weight and gender were available. From seven different studies, the 24-hour profiles of RR- and QT-intervals measured at discrete time points according to one of the seven protocols, were used in this analysis. The mean QT value and the mean RR value from the combined dataset are 390.5 msec (sd. 27.8) and 0.96 msec (sd. 0.15) respectively.

Steps 1 and 2 - QT- and RR-intervals

Using the starting RR-interval values (RR_0) from each of the 776 profiles, a vector of RR-intervals was created to sample from for a new hypothetical population at time 0. Then the median RR_0 was determined. Using the median value, two distributions of RR differences were defined, one each for above and below the median. We chose to use to split the values at the median because it was found that the two distributions statically differ from each other whereas the distributions, when the data is separated by bins of

different sizes or quartiles, were not significantly different. The definition of differences in RR values was defined according to the following formula:

$$\Delta RR1 = M1 - M2; \Delta RR2 = M1 - M3; \ldots, \text{where}$$

M1 = first RR measurement, M2 = second RR measurement, etc …

After the RR value for time 0 for each of the profile in the hypothetical population was sampled, the RR differences for the remaining profile were sampled from one of the two distributions described earlier, depending on the starting value. Finally, a full 24-hour RR profile was back-calculated for each subject based on the differences sampled, according to a typical clinical trial sampling scheme as shown in Table 1. Since the new RR profiles generated were sampled non-parametrically, physiological boundaries were used to prevent unrealistic values.

	Baseline				Dose						Post-dose						
Time	-1	-0.75	-0.5	-0.25	0	0.5	1	1.5	2	2.5	3	4	6	8	12	18	24
PK	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x	x
PD	x		x	x		x	x	x	x			x		x	x		x

Table 1. Sampling scheme for hypothetical trial

QT-intervals for each RR-interval generated according to the aforementioned steps were simulated according to the following relationship,

$$\mathbf{QT = SLP * RR\text{^}EXP + INCPT},$$

between all the pairs of QT and RR values in the healthy subject data, where SLP is the slope parameter of the relationship, EXP is the exponential term quadratic function and INCPT is the parameter of the y-intercept.

Step 3 - New chemical entity

The pharmacokinetic model for the hypothetical compound in step 3 was assumed to be a one-compartment model with typical PK parameter estimates. The mean population estimates used were 0.75 h^{-1} for 1st order absorption constant, Ka, 0.125 h^{-1} for clearance, CL, 100 L for central compartment volume for a hypothetical dose of 100 mg. Each parameter was also fixed with a variability of 30%. The pharmacodynamic effects were

simulated according to the diagram in Figure 2 using a model described in a previous publication [23] with the following relationship:

$$QTc=SLOPE * CONC+QTc0+A * \cos((2 * \pi/24) * (Ctime-phase))$$

Equation 1

The drug-effects (Δy) were fixed at 2 msec for a negative QTc prolongation study, 5 msec for a borderline outcome and 10 msec for a positive effect at the maximum concentration (Cmax). A variability of 1 or 4 msec on top of the mean effect was added to test the sensitivity of the different methodologies. The PD parameters used in the simulations were 6.75 msec for amplitude, 13 h for phase. QTc0 was the actual mean of the QTc baseline values of the simulated population. 30% variability was also applied to all the parameters and a residual error of 5.3 was also included.

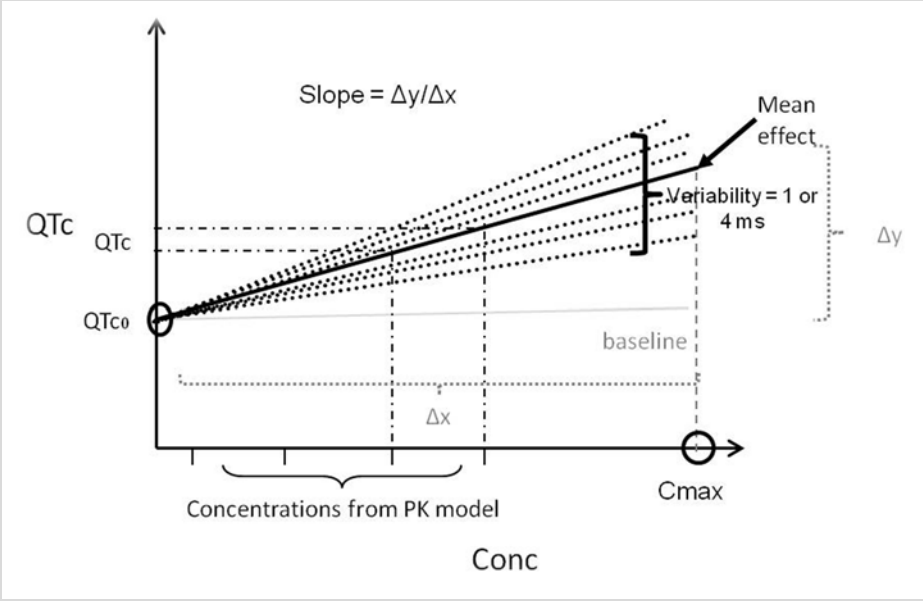


Figure 2. Diagram depicting the relationship between imposed maximum effects at maximum concentration (Cmax) which helps to define the slope values for the PK/PD relationship for the hypothetical compound and thus the resulting drug-induced QTc-interval values.

Step 4 - Positive control

A positive control, such as moxifloxacin, is required to be assessed as part of the thorough-QT study to validate the assay sensitivity of each study [16]. For step 4, the moxifloxacin data was extracted from a two-way crossway, placebo-controlled, single-blind trial aimed at evaluating the effect of repeated oral doses of lamotrigine on cardiac conduction (study SCA104648). Only data from sessions with placebo and moxifloxacin (400mg), which was used as a positive control, were used for the purposes of this study. 137 subjects were included in model building of moxifloxacin.

PHARMACOKINETIC MODELING

The pharmacokinetic data was modeled first using nonlinear mixed effects approach as implemented in NONMEM VI (v 1.0; ICON Development Solutions, Ellicott City, MD) [27]. It was found that an oral administration two-compartment model best described the data with a combination of zero- and first order absorption processes. Model selection was based on the likelihood ratio test, parameter point estimates, and their respective confidence intervals, as well as parameter correlations and goodness-of-fit plots. Given that sample collection for drug concentration quantifications was carried out at time points different from those for ECG measurements, the purpose of the pharmacokinetic model was to allow imputation of drug concentrations at the time point of each ECG measurement so that pair-wise concentration-effect data would be available for pharmacodynamic (PD) modelling.

PHARMACOKINETIC / PHARMACODYNAMIC MODELING AND SIMULATION

Instead of applying traditional maximum likelihood methods to model QTc-interval, a Bayesian adaption of a model was implemented using WinBUGS 1.4 [28, 29] according to Equation 1. The model was implemented in a hierarchical manner in which subject-specific levels (between-subject variability) were considered for all parameters. In depth details of the model building and results were described in an earlier publication [23]. Simulations for the hypothetical cohort were carried out according the final model with zero-order absorption value of 2.21 h^{-1}, first-order absorption value of 0.629 h^{-1}, 13.4 L/h for clearance, 122 L as the volume for central

compartment, 55.4 L for the peripheral compartment volume, 78.4 h^{-1} as the inter-compartmental clearance and 0.004 msec/ng/mL as the slope value for the PKPD relationship.

STUDY DESIGN SCENARIOS

Three different clinical study designs (crossover with full day baseline, crossover with 3 pre-dose baseline and parallel), number of subjects (16, 30, 46, 60, 90 and 120) and effect sizes of the hypothetical compound (2, 5 and 10 msec) were tested. A total of 36 scenarios were created, Table 2 shows the details of permutation.

Study Design	Number of Subjects	Effect Size
Crossover with time-matched (full day) baseline / Crossover with 3 predose baseline	16	2
		5
		10
	30	2
		5
		10
	46	2
		5
		10
	60	2
		5
		10
Parallel	30	2
		5
		10
	60	2
		5
		10
	90	2
		5
		10
	120	2
		5
		10

Table 2. Study designs for hypothetical trials

For each of the thirty-six scenarios, two levels of between subject-variability on the slope parameter were considered, namely, 1 and 4 msec. Finally, one hundred trials were generated for each of the scenarios.

DATASET SIMULATIONS AND RESULTS ANALYSIS

Using R2.8 (The R project - http://www.r-project.org/) [30], the entire simulation exercise was implemented as stated above. It was also used to control the execution and collection of results from other software, namely NONMEM and WinBUGS. After the fully completed datasets were generated, the software was used again to analyse the results according to the criteria described below.

SENSITIVITY AND SPECIFICITY OF THE DOUBLE-DELTA METHOD AND HIERARCHICAL MODELLING APPROACH

Each dataset was analysed for its propensity for QTc-interval prolongation by the double-delta method, as proposed in the ICH E14 guideline [1] and by the Bayesian hierarchical modelling approach [23] where all the parameters in the model were estimated. All the results were tabulated and compared to the true drug-effects imposed. For each combination of trial design and number of subjects, specificity and sensitivity were calculated according the formulas:

$$\text{specificity} = \frac{\text{number of True Negatives}}{\text{number of True Negatives} + \text{number of False Positives}}$$

$$\text{sensitivity} = \frac{\text{number of True Positives}}{\text{number of True Positives} + \text{number of False Negatives}}$$

We defined false positive trials with drug effects of 2 or 5 msec where the resulting analysis indicates a QTc-interval prolongation of greater or equal to

10 msec. On the other hand, false negatives were based on analysis done with drug-effects equal to 10 msec. Should the final results from respective methods do not detect a prolongation of greater or equal to 10 msec, then the trial would be deemed a false negative.

10.3 RESULTS

With 36 permutations for design scenarios, 100 simulated trials each, and both 1 msec and 4 msec variability imposed on the slope value of the hypothetical drug in the PKPD relationship, a total of 7200 trials were simulated. An example of the pharmacokinetic profiles of the new chemical entity is displayed in Figure 3.

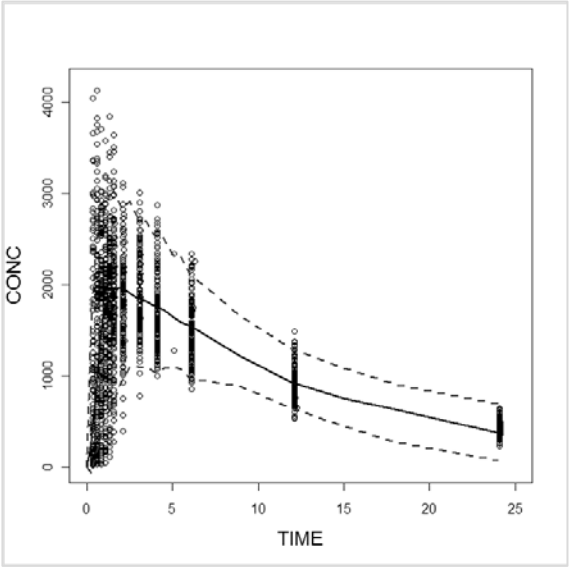


Figure 3. An example of the pharmacokinetic profiles over time for the hypothetical compound used for the purpose of the simulations.

4 msec variability on slope

A tabulation of the final comparison for the sensitivity and specificity between the traditional double-delta method and the Bayesian method is shown in Table 3a. The table shows the outcomes for the different trial designs tested, namely crossover studies with time-matched full day baseline

measurements (CRtm), crossover design with 3 pre-dose baseline measurements (CRbl) and parallel study design (PA), along with various numbers of subjects per trial. The false negative and positive percentages over different number of subjects per trial design for each design are shown in Figure 4 for CRtm, CRbl and PA designs.

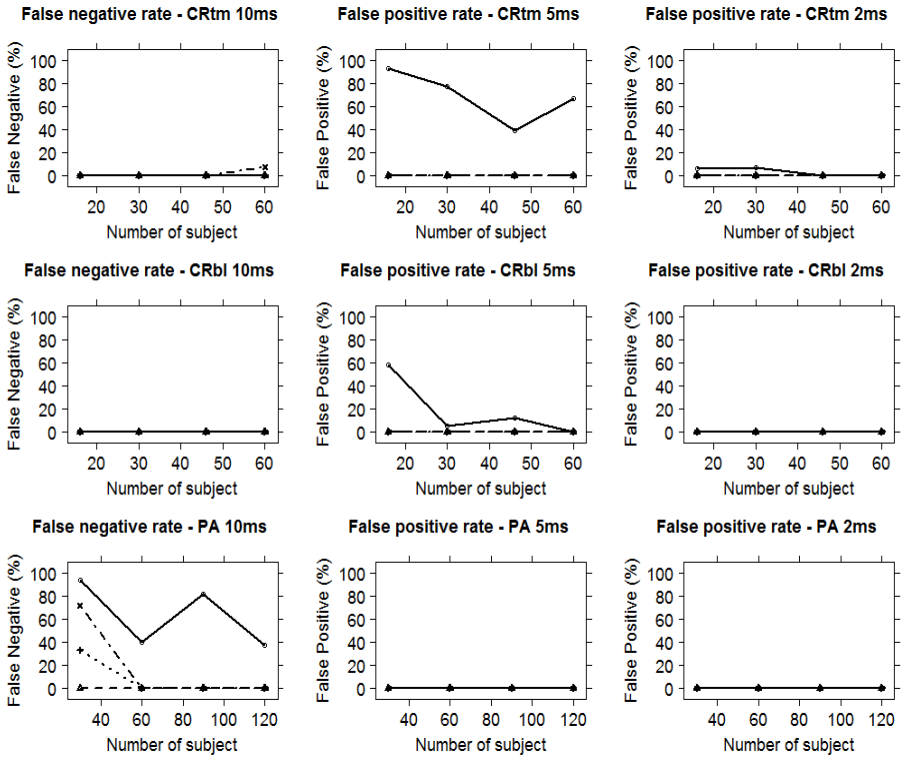


Figure 4. False negative and positive percentages over different number of subject per trial design (10msec, 5msec, 2msec) for the crossover time-matched full day baseline measurements (top panels), crossover 3 pre-dose baseline measurements (middle panels), and parallel study design scenario (bottom panels). Solid line denotes result from double delta method, Δ denotes a trial is positive when at least 99% of all BUGS simulations are positive, + denotes a trial is positive when at least 95% of all BUGS simulations are positive, × denotes a trial is positive when at least 90% of all BUGS simulations are positive.

1 msec variability on slope

When 1 msec variability was applied on the slope parameter for the PKPD relationship of the hypothetical compound, the trend of the results were

similar to the 4 msec variability. The tabulation showing the results for specificity and sensitivity is displayed in Table 3b.

4 msec var on SLP		CRtm 16	CRtm 30	CRtm 46	CRtm 60	CRbl 16	CRbl 30	CRbl 46	CRbl 60	PA 30	PA 60	PA 90	PA 120
DD	*Specificity*	0.505	0.58	0.805	0.665	0.71	0.965	0.94	1	1	1	1	1
	Sensitivity	1	1	1	1	1	1	1	1	0.06	0.6	0.18	0.6
BUGS	*Specificity*	1	1	1	1	1	1	1	1	1	1	1	1
	Sensitivity	1	1	1	1	1	1	1	1	1	1	1	1

CRtm = crossover with time-matched baseline, CRbl = crossover with 3 pre-dose baseline, PA = parallel design

Table 3a. Specificity and sensitivity tabulation for the 3600 simulations with 4 msec variability applied to the slope parameter comparing the double delta method (DD) and the Bayesian approach (BUGS).

1 ms var on SLP		CRtm 16	CRtm 30	CRtm 46	CRtm 60	CRbl 16	CRbl 30	CRbl 46	CRbl 60	PA 30	PA 60	PA 90	PA 120
	Specificity	0.685	0.79	0.85	0.855	0.97	0.97	0.99	1	1	1	1	1
	Sensitivity	1	1	1	1	1	1	1	1	0.65	0.41	0.24	0.14
DD	*Specificity*	1	1	1	1	1	1	1	1	1	1	1	1
BUGS	*Sensitivity*	1	1	1	1	1	1	1	1	1	1	1	1

CRtm = crossover with time-matched baseline, CRbl = crossover with 3 pre-dose baseline, PA = parallel design

Table 3b. Specificity and sensitivity tabulation for the 3600 simulations with 1 msec variability applied to the slope parameter comparing the double delta method (DD) and the Bayesian approach (BUGS).

10.4 DISCUSSIONS

In summary, the Bayesian approach has higher or equal specificity and sensitivity compared to the double-delta method in all trial designs. False positive trials are only detected with the double-delta method where lower specificity is observed with smaller number of subjects per trial. Lower specificity is also seen with full day baseline measurements as compared to only 3 pre-dose baseline measurements. In addition, false negative trials are more frequently detected with parallel design. Lower sensitivity is achieved with the double-delta method with lesser number of subjects per trial as expected.

The results from this study highly suggest that the Bayesian approach should be considered when analysing the propensity of drug-induced QTc-interval prolongation when determining the proarrhythmic-genicity of new compounds in the future. The novel approach not only has the ability to correctly identify the effect size at the therapeutic exposure range [24] due to its nature to describe the PKPD relationship of the compound [25, 26], but it is also sensitive enough to perform the analysis with smaller number of subjects and has the flexibility to accommodate different trial designs.

Keeping in mind that the high cost of studies is mainly attributed to the requirement of large trials to satisfy the power needed to show an effect of a QT-interval prolongation of 5 msec, being able to achieve the same results with lower sample sizes would minimise the burden. In the present analysis, very specific conditions and properties in pharmacokinetic and pharmacodynamic have been placed in the hypothetical compound. However, the Bayesian approach is not limited to those specific conditions and can be used towards a variety of compounds, such as those where covariate effects alter the PK, compounds with concurrent effect on RR as well as on QT or other cardiovascular parameter. The same flexibility can be applied to manually-read or machine-read ECG measurements or any other algorithms, as well as other methods used to resolve between-observer variability in the observations, as those method are independent to the way ECG is read.

Another advantage of the PKPD model is that it can handle multiplicity issues arising from having to establish non-inferiority at each time point. The QTc prolongation is established continuously at the range of observable

concentration, rather than at each discrete measurement. Conversely, to minimise the problem, investigators may choose to take measurements only at a bare minimum required number of time points when using the double-delta approach.

Ultimately, the newly proposed Bayesian method still depends on the placement of PK and PD samples taken for the study. If the sampling schedule misses the maximum concentration that leads to the true maximum effect being absent from the data or Emax is unrecorded independently, then the analysis would not be able to capture the true proarrhythmic property of the compound. However, should the nature of the PKPD relationship is known in advance, such as hysteresis, adjustments can be made to the Bayesian model to accommodate for the use of an effect compartment.

In conclusion, we have conducted a simulation of a hypothetical population on a compound with imposed drug effects so that we can clearly compare the results of the drug-induced QTc-interval prolongation analysis using a previously published Bayesian approach with the double-delta method as proposed by the ICH E14 guideline. The outcomes have shown that the Bayesian approach is able to detect the prolongation of QTc-interval with equal or better accuracy and precision at various effect sizes, using different number of subjects and with a variety of trial designs. These encouraging facts are envisioned to provide an alternative to the current methodology used in TQT studies as well as a strategy to minimise the cost of drug development.

REFERENCES

[1] Guidance for Industry: ICH E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrythmic Potential for Non-Antiarrythmic Drugs.U.S. Department of Health and Human Services. Food and Drug Administration. (2005)

[2] Hutmacher MM, Chapel S, Agin MA, Fleishaker JC, Lalonde RL. Performance characteristics for some typical QT study designs under the ICH E-14 guidance. J Clin Pharmacol 2008 Feb;48(2):215-24.

[3] Malik M, Hnatkova K, Batchvarov V, Gang Y, Smetana P, Camm AJ. Sample size, power calculations, and their implications for the cost of thorough studies of drug induced QT interval prolongation. Pacing Clin Electrophysiol 2004 Dec;27(12):1659-69.

[4] Natekar M, Mahajan V, Satra A, O'Kelly M, Karnad DR. Use of ANOVA to estimate inter- and intra-reader variability for a group of readers in thorough QT/QTc studies. Clin Pharmacol Ther 2008 Mar;83(3):489-91.

[5] Salvi V, Karnad DR, Panicker GK, Kothari S. Update on the evaluation of a new drug for effects on cardiac repolarization in humans: issues in early drug development. Br J Pharmacol 2010 Jan;159(1):34-48.

[6] Darpo B, Nebout T, Sager PT. Clinical evaluation of QT/QTc prolongation and proarrhythmic potential for nonantiarrhythmic drugs: the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use E14 guideline. J Clin Pharmacol 2006 May;46(5):498-507.

[7] Sethuraman V, Sun Q. Impact of baseline ECG collection on the planning, analysis and interpretation of 'thorough' QT trials. Pharm Stat 2009 Apr-Jun;8(2):113-24.

[8] Tian H, Natarajan J. Effect of baseline measurement on the change from baseline in QTc intervals. J Biopharm Stat 2008;18(3):542-52.

[9] Bloomfield DM, Kost JT, Ghosh K, Hreniuk D, Hickey LA, Guitierrez MJ, et al. The effect of moxifloxacin on QTc and implications for the design of thorough QT studies. Clin Pharmacol Ther 2008 Oct;84(4):475-80.

[10] Ma H, Smith B, Dmitrienko A. Statistical analysis methods for QT/QTc prolongation. J Biopharm Stat 2008;18(3):553-63.

[11] Hnatkova K, Malik M. "Optimum" formulae for heart rate correction of the QT interval. Pacing Clin Electrophysiol 1999 Nov;22(11):1683-7.

[12] Malik M, Farbom P, Batchvarov V, Hnatkova K, Camm AJ. Relation between QT and RR intervals is highly individual among healthy subjects: implications for heart rate correction of the QT interval. Heart 2002 Mar;87(3):220-8.

[13] Boos DD, Hoffman D, Kringle R, Zhang J. New confidence bounds for QT studies. Stat Med 2007 Sep 10;26(20):3801-17.

[14] Hosmane B, Locke C, Chiu YL. Sample size and power estimation in "thorough" QT/QTc studies with parallel group design. J Biopharm Stat 2010 May;20(3):578-86.

[15] Anand SP, Murray SC, Koch GG. Sample size calculations for crossover thorough QT studies: satisfaction of regulatory threshold and assay sensitivity. J Biopharm Stat 2010 May;20(3):587-603.

[16] Stylianou A, Roger J, Stephens K. A statistical assessment of QT data following placebo and moxifloxacin dosing in thorough QT studies. J Biopharm Stat 2008;18(3):502-16.

[17] Couderc JP, Garnett C, Li M, Handzel R, McNitt S, Xia X, et al. Highly automated QT measurement techniques in 7 thorough QT studies implemented under ICH E14 guidelines. Ann Noninvasive Electrocardiol 2011 Jan;16(1):13-24.

[18] Baranowski R, Poplawska W, Buchner T, Chojnowska L, Rydlewska-Sadowska W. Day-to-day reproducibility of Holter beat-by-beat analysis of repolarisation. Acta Cardiol 2003 Jun;58(3):185-9.

[19] De Groote K, Suys B, Deleeck A, De Wolf D, Matthys D, Van Overmeire B. How accurately can QT interval be measured in newborn infants? Eur J Pediatr 2003 Dec;162(12):875-9.

[20] Kautzner J, Yi G, Camm AJ, Malik M. Short- and long-term reproducibility of QT, QTc, and QT dispersion measurement in healthy subjects. Pacing Clin Electrophysiol 1994 May;17(5 Pt 1):928-37.

[21] Morganroth J. Cardiac repolarization and the safety of new drugs defined by electrocardiography. Clin Pharmacol Ther 2007 Jan;81(1):108-13.

[22] Chapel S, Hutmacher MM, Bockbrader H, de Greef R, Lalonde RL. 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 2011 Jan;89(1):75-80.

[23] Chain AS, Krudys KM, Danhof M, Della Pasqua O. Assessing the Probability of Drug-Induced QTc-Interval Prolongation During Clinical Drug Development. Clin Pharmacol Ther 2011 Dec;90(6):867-75.

[24] Chan PL, Holford NH. Drug treatment effects on disease progression. Annu Rev Pharmacol Toxicol 2001;41:625-59.

[25] Garnett CE, Beasley N, Bhattaram VA, Jadhav PR, Madabushi R, Stockbridge N, et al. Concentration-QT relationships play a key role in the evaluation of proarrhythmic risk during regulatory review. J Clin Pharmacol 2008 Jan;48(1):13-8.

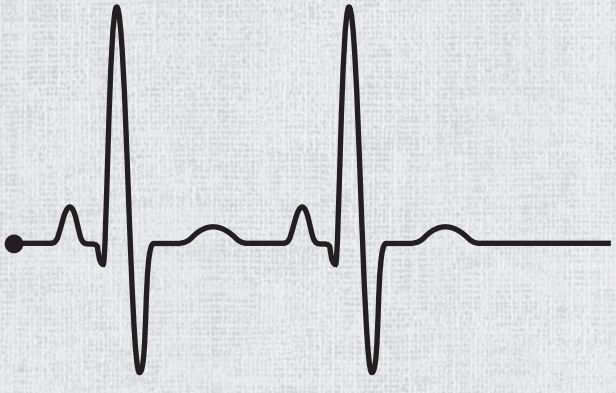
[26] Russell T, Riley SP, Cook JA, Lalonde RL. A perspective on the use of concentration-QT modeling in drug development. J Clin Pharmacol 2008 Jan;48(1):9-12.

[27] Beal S, Sheiner, L. NONMEM user's guide NONMEM project group. San Francisco; 1999.

[28] Lunn DJ, Thomas, A., Best, N., Spiegelhalter, D. WinBUGS - A Bayesian modelling framework: concepts, structure, and extensibility. Stat Comput 2000;10:325-37.

[29] Spiegelhalter D, Thomas, A., Best, N. & Gilks W. BUGS: Bayesian Inference Using Gibbs Sampling, Version 1.4.1 Cambridge, UK: MRC Biostatistics Unit.

[30] R Development Core Team, R: a language and environment for statistical compting. In: computing Rffs, editor. Vienna, Austria; 2010.

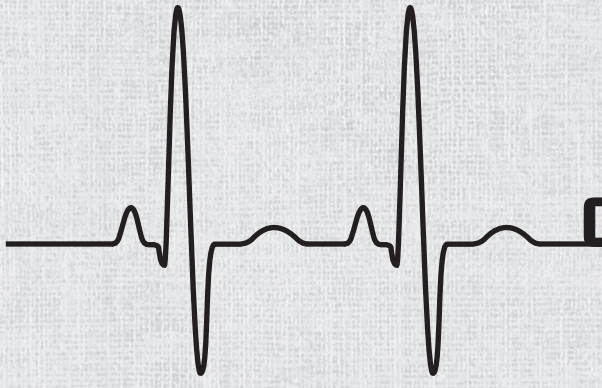


Can First-Time-In-Human Trials Replace Thorough QT Studies?

ASY Chain (1), F Bellanti (1), M Danhof (1),
MCJM Sturkenboom (2, 3), OE Della Pasqua (1,4)

(1) Leiden/Amsterdam Centre for Drug Research,
Division of Pharmacology, LeidenUniversity, Leiden, The Netherlands
(2) Department of Medical Informatics and Epidemiology,
Erasmus Medical Centre, The Netherlands
(3) Department of Epidemiology, Erasmus Medical Centre,
The Netherlands
(4) Clinical Pharmacology and Discovery Medicine,
GlaxoSmithKline, Stockley Park, Uxbridge, United Kingdom

Submitted for publication



Chapter 11

ABSTRACT

The aim of the present study was to investigate if the first-time-in-human (FTIH) trials can be modified and optimised to achieve the same information as thorough-QT (TQT) studies. Three different FTIH trial designs with various sample sizes were simulated and analysed using negative and positive hypothetical drug-effects. First, a traditional dose escalation study design was simulated and analysed according to a standard protocol. Second, an additional dose of moxifloxacin was incorporated into the traditional study design where only PD measurements were sampled and historical PK information was used. A similar scenario was created last where both PK and PD observations are simulated. All scenarios performed, adequately captured the drug-induced QT-interval prolongation with the probabilities correctly indicating the likelihood of a change greater than or equal to 10 msec. By modifying the traditional FTIH studies, we were able to gain valuable safety information early in development which can facilitate making decisions.

11.1 INTRODUCTION

Since 2005, the International Conference on Harmonisation (ICH) introduced the topic E14 guideline mandating the systematic evaluation of QT prolongation [2]. Regulatory agencies expect evidence on the propensity of a new chemical / biological entity to cause pro-arrhythmia, as determined by data arising from a thorough-QT (TQT) study. Numerous arguments have been proposed to defend the need for a separate clinical trial to address this specific safety concern [3]. Among other factors, the requirement for manually read ECGs, a positive control treatment arm and time-matched sampling scheme have been used as arguments to exclude alternative options for evidence on cardiovascular safety (e.g. aggregated summary of safety or meta-analysis).

Oncoming evidence clearly show however, that the current recommendations for a TQT study appear to have been reached without careful evaluation of pharmacological, statistical and technical aspects underlying the assessment of drug-induced effect on cardiac repolarisation. From a pharmacological perspective, TQT studies neglect the underlying concentration-effect relationship, which ultimately determines the clinical relevance of drug-induced effects [1, 4]. Statistically, study design and sampling schemes overlook type I and II error rates, in particular when replicates or more frequent sampling are used [5, 6]. Furthermore, recent data suggest that despite the expenditure of 2-7 million US dollars, conclusions about the magnitude of drug effects are possibly biased [7]. From a technical perspective, the introduction of manually read tracings has not eliminated all the concerns regarding measurement bias. Surface ECG algorithms have intrinsic limitations which cannot be fully mitigated by manual reading procedures. More recently, the value of current regulations has been questioned even from a cost-effectiveness perspective [7, 8].

Often opposition to the requirement for a mandatory TQT study is perceived as an attempt to evade the need to safeguard patients. For those less familiar with the evolution of the proposal that resulted in the current guidelines, it should be clear that from a clinical and scientific perspective, there is no dispute about the importance of obtaining accurate evidence of cardiovascular safety. Undoubtedly, what one must continue to question is how such evidence is generated, analysed and interpreted.

Given the statistical flaws, the scientific oversight, the ethical burden and financial consequences of study design implementation recommended by ICH E14, the current investigation assesses the feasibility of using first-time-in-human (FTIH) studies as the basis for evidence synthesis to investigate the propensity for pro-arrhythmic effects. FTIH studies are a mandatory step in the drug development process. Traditionally, the studies are small, with single and multiple escalated doses, normally reaching supra-therapeutic levels in healthy subjects or in the target patient population. The purpose of such escalating doses in a FTIH study is to characterise pharmacokinetics (ADME - absorption, distribution, metabolism, excretion properties), pharmacodynamics, safety and tolerability. In principle, the doses or dose range evaluated during escalation should enable the evaluation of the concentration-effect curve, providing evidence for drug effects not only at therapeutic level, but also supra-therapeutic levels. From a safety and tolerability perspective, pharmacodynamic measures are monitored frequently or continuously throughout the dosing interval in parallel to pharmacokinetic sampling. Specifically, 12-lead ECG or Holters are used for cardiovascular safety monitoring according to the common standards applied also for other phase I clinical trials [9].

Recently, using a model-based approach, we have shown the advantages of using a Bayesian hierarchical model to quantify drug-induced effects on QTc-interval [1]. Our approach also provides estimates of the probability of QTc prolongation *versus* a wide range of drug concentrations, providing immediate translation of the findings in terms of their clinical relevance. It should be noted that the drug-specific parameter in this model (slope) does not independently indicate whether there is a risk associated with the use of the compound. Rather, it provides an unambiguous description of the PKPD relationship, taking into account various sources of variability. We anticipate therefore that the use of an integrated hierarchical model for the assessment of the propensity for QTc-interval prolongation will be sufficiently robust to overcome some of the criticisms around the use of different study designs and the reliance on automated ECG measurements.

We also challenge the requirement for the continuous assessment of individual sensitivity to QTc-interval prolongation. In fact, it has been demonstrated that other interventions or procedures such as food and liquid intake and exercise may all be used as proxy for sensitivity [10]. Yet, we do not exclude the possibility of including a benchmark or positive control arm

in a typical FTIH trial, if necessary. The many historical studies with moxifloxacin [11] can be used in an integrated manner as benchmark or as priors during data analysis.

11.2 METHODS

SIMULATION OF EXPERIMENTAL VARIABLES IN A HYPOTHETICAL POPULATION

Assumptions

The steps taken to generate a dataset for a hypothetical population were not done without certain assumptions. For instance, we assumed that the random noise in the simulated data reflects the actual variability in measurements. In manually read ECG data, there exist between-observer variability and potentially other sources of bias due to algorithm used to harmonise the differences, while machine-read ECG values also includes measurement errors. We have also assumed there are no systematic biases due to automated ECG readings that yield reliable measurements with constant variance throughout the simulation.

No demographic or other covariates were simulated for the hypothetical population because these were deemed as random factors and as such will not interfere with the drug-induced effect size estimation. Furthermore, we made the assumption that there was no metabolites formation which could lead to the delayed onset of effects. The parent drug solely affected QT-intervals and not on RR, thus only minor within-subject variability existed in the QT/RR relationship.

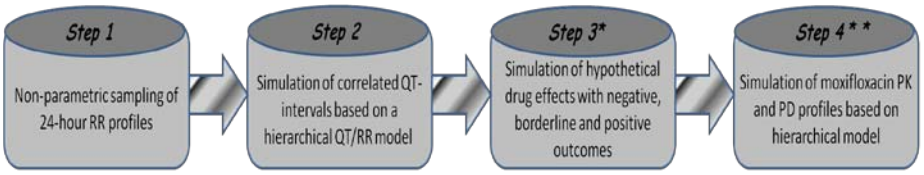
Data generation

An overview to generate a dataset for a hypothetical population to compare the methods for analysing QTc-interval prolongation is shown in Figure 1.

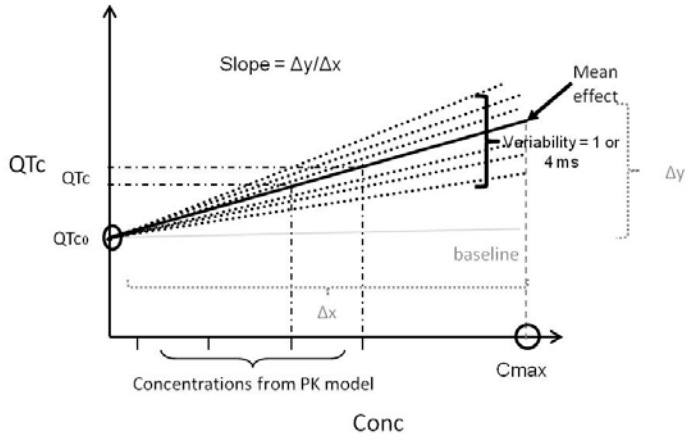
Study population demographics

The data for this study was derived from a large pool of healthy volunteer ECG data, totaling 776 subjects (339 males and 437 females), with mean age of 31.1 years (sd. 9.4). Variables such as QT, QTc, RR, VR, QRS, PR, age,

weight and gender were available. From seven different studies, the 24-hour profiles of RR- and QT-intervals measured at discrete time points according to one of the seven protocols, were used in this analysis. The mean QT value and the mean RR value from the combined dataset are 390.5 msec (sd. 27.8) and 0.96 msec (sd. 0.15) respectively.



*



**

	Baseline			Dose		Post-dose											
Time	-1	-0.75	-0.5	-0.25	0	0.5	1	1.5	2	2.5	3	4	6	8	12	18	24
PK	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x	x
PD	x		x	x		x	x	x	x			x		x	x		x

*Figure 1. Flow diagram describing the steps taken to generate a full dataset for a hypothetical population. * A closer examination on the relationship between imposed maximum effects at maximum concentration (Cmax) which helps to define the slope values for the PK/PD relationship for the hypothetical compound and thus the resulting drug-induced QTc-interval values. ** The typical sampling scheme for a clinical trial used in step 4.*

Steps 1 and 2 - QT- and RR-intervals

Using the starting RR-interval values (RR_0) from each of the 776 profiles, a vector of RR-intervals was created to sample from for a new hypothetical population at time 0. Then the median RR_0 was determined. Using the median value, two distributions of RR differences were defined, one each for above and below the median. We chose to use to split the values at the median because it was found that the two distributions statically differ from each other whereas the distributions, when the data is separate by bins of different sizes or quartiles, were not significantly different. The definition of differences in RR values was defined according to the following formula:

$$\mathbf{\Delta RR1 = M1 - M2; \Delta RR2 = M1 - M3}; \ldots, \text{where}$$

M1 = first RR measurement, M2 = second RR measurement, etc …

After the RR value for time 0 for each of the profile in the hypothetical population was sampled, the RR differences for the remaining profile were sampled from one of the two distributions described earlier, depending on the starting value. Finally, a full 24-hour RR profile was back-calculated for each subject based on the differences sampled, according to a typical clinical trial sampling scheme as part of Figure 1. Since the new RR profiles generated were sampled non-parametrically, physiological boundaries were used to prevent unrealistic values.

QT-intervals for each RR-interval generated according to the aforementioned steps were simulated according to the following relationship,

$$\mathbf{QT = SLP * RR\text{^}EXP + INCPT},$$

between all the pairs of QT and RR values in the healthy subject data, where SLP is the slope parameter of the relationship, EXP is the exponential term quadratic function and INCPT is the parameter of the y-intercept.

Step 3 - New chemical entity

The pharmacokinetic (PK) model for the hypothetical compound in step 3 was assumed to be a one-compartment model with typical PK parameter estimates. The mean population estimates used were 0.75 h^{-1} for 1st order absorption constant, Ka, 0.125 h^{-1} for clearance, CL, 100 L for central compartment volume for a hypothetical dose of 100 mg. From the initial PK model, concentration values for doses of 5 mg, 10 mg, 25 mg, 50 mg, 250mg and 500 mg were simulated. Each parameter was also fixed with a variability of 30%. The pharmacodynamic effects were simulated according to the sampling scheme shown as part of Figure 1 and the model described in a previous publication [1] with the following relationship:

$$QTc=SLOPE * CONC+QTc0+A * \cos((2 * \pi/24) * (Ctime-phase))$$

Equation 1

The drug-effects (Δy) were fixed at 2 msec for a negative QTc prolongation study, 5 msec for a borderline outcome and 10 msec for a positive effect at the maximum concentration (Cmax). A variability of 1 or 4 msec on top of the mean effect was added to test the sensitivity of the different methodologies. The PD parameters used in the simulations were 6.75 msec for amplitude, 13 h for phase. QTc0 was the actual mean of the QTc baseline values of the simulated population. 30% variability was also applied to all the parameters and a residual error of 5.3 was also included.

Step 4 - Positive control

For step 4, the moxifloxacin data was extracted from a two-way crossway, placebo-controlled, single-blind trial aimed at evaluating the effect of repeated oral doses of lamotrigine on cardiac conduction (study SCA104648). Only data from sessions with placebo and moxifloxacin (400mg), which was used as a positive control, were used for the purposes of this study. 137 subjects were included in model building of moxifloxacin.

PHARMACOKINETIC MODELLING

The pharmacokinetic data was modelled first using nonlinear mixed effects approach as implemented in NONMEM VI (v 1.0; ICON Development

Solutions, Ellicott City, MD) [12]. It was found that an oral administration two-compartment model best described the data with a combination of zero- and first order absorption processes. Model selection was based on the likelihood ratio test, parameter point estimates, and their respective confidence intervals, as well as parameter correlations and goodness-of-fit plots. Given that sample collection for drug concentration quantifications was carried out at time points different from those for ECG measurements, the purpose of the pharmacokinetics model was to allow imputation of drug concentrations at the time point of each ECG measurement so that pair-wise concentration-effect data would be available for pharmacodynamic (PD) modelling.

PHARMACOKINETIC / PHARMACODYNAMIC MODELLING AND SIMULATION

Instead of applying traditional maximum likelihood methods to model QTc-interval, a Bayesian adaption of a model was implemented using WinBUGS 1.4 [13, 14] according to Equation 1. The model was implemented in a hierarchical manner in which subject-specific levels (between-subject variability) were considered for all parameters. In depth details of the model building and results were described in an earlier publication [1]. Simulations for the hypothetical cohort were carried out according the final model with zero-order absorption value of 2.21 h^{-1}, first-order absorption value of 0.629 h^{-1}, 13.4 L/h for clearance, 122 L as the volume for central compartment, 55.4 L for the peripheral compartment volume, 78.4 h^{-1} as the inter-compartmental clearance and 0.004 msec/ng/mL as the slope value for the PKPD relationship.

STUDY DESIGN SCENARIOS

Three different FTIH trial designs, number of subjects (12, 18, 27) and drug-effect sizes of the hypothetical compound were tested, as shown in Table 1. Randomisation schemes of each dose-escalation studies are depicted in Tables 2a-c.

In total, 150 simulations were performed for each scenario. The hypothetical compounds were assumed to have no effect (2 msec), or significant effect (10 msec) on QTc-interval. Doses of 5, 10, 25, 50, 100, 250 and 500 mg of hypothetical compounds were given in all scenarios. Pharmacokinetic (PK)

and pharmacodynamic (PD) sampling schemes were simulated according to standard FTIH protocol. Data analysis was performed in WinBUGS v 1.4.3, where the probabilities of QTc-interval prolongation were derived from the underlying QTc-interval *versus* concentration relationship [1].

	Scenario 1 (Typical FTIH)	Scenario 2 (Modified 1 FTIH)	Scenario 3 (Modified 2 FTIH)
Trial Design	- 5 sessions - Placebo-controlled trial - 4 consecutively escalating doses of hypothetical compound	- 6 sessions - Placebo-controlled trial - 4 consecutively escalating doses of hypothetical compound - Single dose of moxifloxacin	- 5 sessions - Placebo-controlled trial - 3 consecutively escalating doses of hypothetical compound - Single dose of moxifloxacin
Sample Size	N = 12	N = 12 and 18	N = 18 and 27
Randomisation	Two cohorts: - Cohort 1: 5, 10, 25 and 50 mg of active drug - Cohort 2: 50, 100, 250 and 500 mg of active drug	Two cohorts: - Cohort 1: 5, 10, 25, and 50 mg of active drug - Cohort 2: 50, 100, 250 and 500 mg of active drug - All received moxifloxacin	Three cohorts: - Cohort 1: 5, 10 and 25 mg of active drug - Cohort 2: 25, 50 and 100 mg of active drug - Cohort 3: 100, 250 and 500 mg of active drug - All received moxifloxacin
Additional Information	Trial was analysed with uninformative priors as well as using historical moxifloxacin parameter estimates as informative prior	Only ECG were taken when moxifloxacin was administered. Historical population PK values were used for both cohorts.	PK and ECG were taken for all five sessions (including moxifloxacin)

Table 1. Trial design specifications are described for each of the different FTIH studies.

DATASET SIMULATIONS AND RESULTS ANALYSIS

Using R2.8 (The R project - http://www.r-project.org/) [15], the entire simulation exercise was implemented as stated above. It was also used to

control the execution and collection of results from other software, namely NONMEM and WinBUGS. After the fully completed datasets were generated, the software was used again to analyse the results according to the criteria described in the next section.

SUBJ	DAY 1	DAY 8	DAY 15	DAY 21	DAY 28
1	PLACEBO	D1	D2	D3	D4
2	D1	D2	PLACEBO	D3	D4
3	D1	PLACEBO	D2	D3	D4
4	D1	D2	D3	D4	PLACEBO
5	D1	D2	D3	PLACEBO	D4
6	D1	D2	PLACEBO	D3	D4

SUBJ	DAY 1	DAY 8	DAY 15	DAY 21	DAY 28	DAY 35
1	PLACEBO	D1	D2	D3	D4	Moxi
2	D1	D2	PLACEBO	D3	D4	Moxi
3	D1	PLACEBO	D2	D3	D4	Moxi
4	D1	D2	D3	D4	PLACEBO	Moxi
5	D1	D2	D3	PLACEBO	D4	Moxi
6	D1	D2	PLACEBO	D3	D4	Moxi

SUBJ	DAY 1	DAY 8	DAY 15	DAY 21	DAY 28
1	PLACEBO	D1	D2	D3	Moxi
2	D1	D2	PLACEBO	D3	Moxi
3	D1	PLACEBO	D2	D3	Moxi
4	D1	D2	PLACEBO	D3	Moxi
5	D1	PLACEBO	D2	D3	Moxi
6	D1	D2	D3	PLACEBO	Moxi
7	PLACEBO	D1	D2	D3	Moxi
8	D1	D2	D3	PLACEBO	Moxi
9	D1	D2	PLACEBO	D3	Moxi

Tables 2a-c. Detailed randomisation scheme for each design: traditional FTIH (top), modified 1 FTIH (middle), and modified 2 FTIH (bottom).

SENSITIVITY AND SPECIFICITY OF THE DOUBLE-DELTA METHOD AND HIERARCHICAL MODELLING APPROACH

Each dataset was analysed for its propensity for QTc-interval prolongation by the double-delta method, as proposed in the ICH E14 guideline [2] and by a Bayesian hierarchical modelling approach [1] where all the parameters in the model were estimated. All the results were tabulated and compared to the true drug-effects imposed. For each combination of trial design and number of subjects, specificity and sensitivity were calculated according the formulas:

$$\text{specificity} = \frac{\text{number of True Negatives}}{\text{number of True Negatives} + \text{number of False Positives}}$$

$$\text{sensitivity} = \frac{\text{number of True Positives}}{\text{number of True Positives} + \text{number of False Negatives}}$$

We defined false positive trials with drug effects of 2 or 5 msec where the resulting analysis indicates a QTc-interval prolongation of greater or equal to 10 msec. On the other hand, false negatives were based on analysis done with drug-effects equal to 10 msec. Should the final results from respective methods do not detect a prolongation of greater or equal to 10 msec, then the trial would be deemed a false negative.

11.3 RESULTS

The three different scenarios required a total 7200 FTIH datasets to be simulated. The top panels of Figure 2 represent an example of the dose escalation PK and PD profiles of the hypothetical compound. Moxifloxacin (positive control) profiles are shown in the bottom panels of Figures 2.

It was possible to accurately predict the risk of 10 msec QT prolongation using simulated FTIH data, with and without evidence from a positive control treatment arm. Figures 3 and 4 present a selection of the most relevant findings. The uncertainty in the relationship between QT-interval prolongation versus concentration and the probability of an increase of 10 msec was larger with smaller population size in all cases.

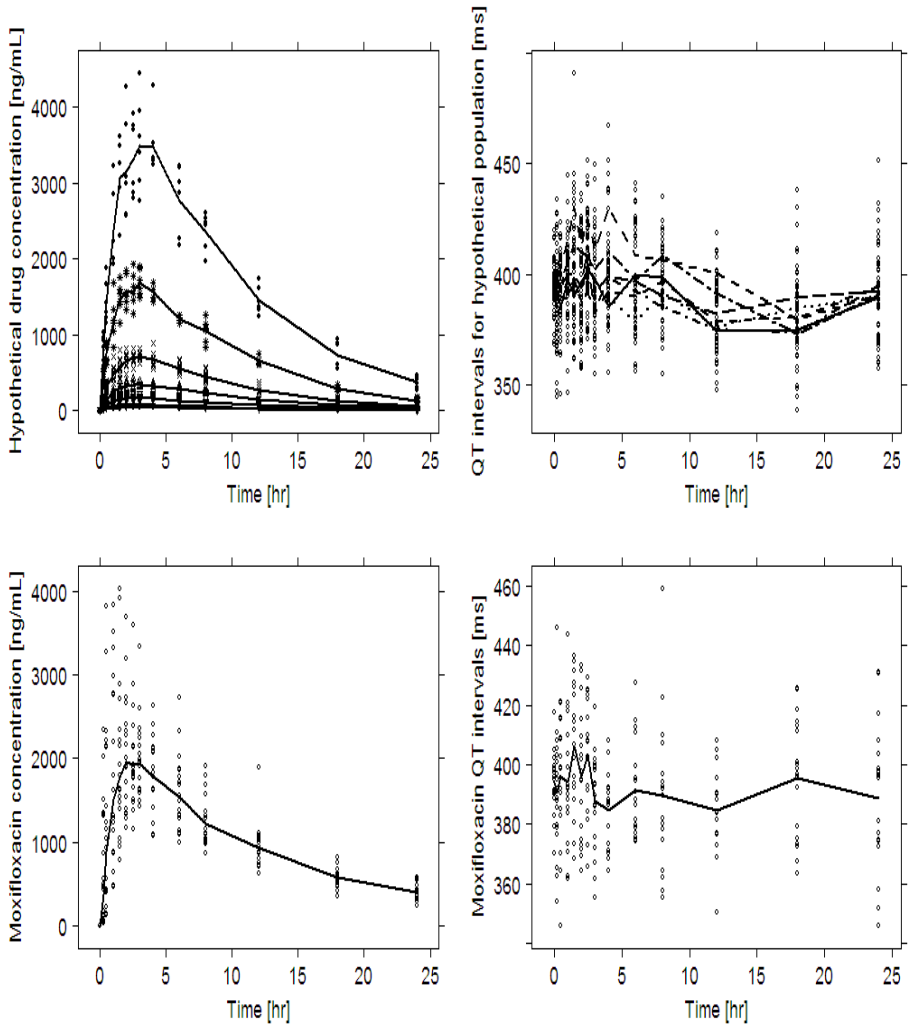


Figure 2. Typical PK (left), PD (right) profiles of the hypothetical compound (top panels), PK (left) and PD (right) profiles of moxifloxacin (lower panels) used in the simulated FTIH studies.

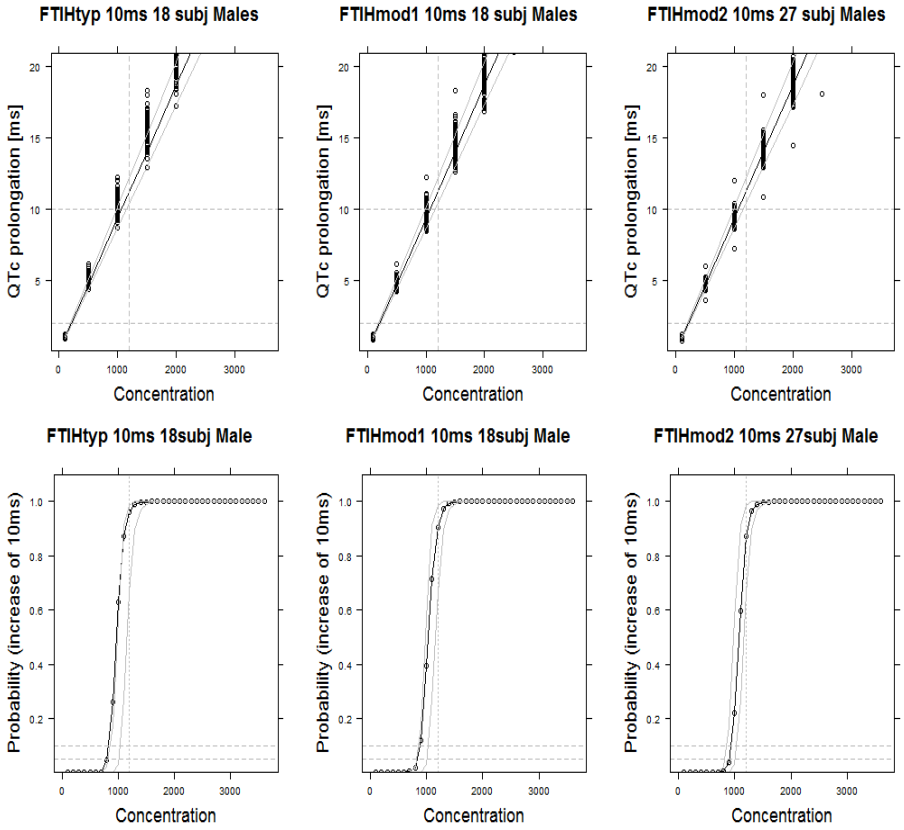


Figure 3. Relationship between QT prolongation (top panels) and probability of ≥10 msec increase (bottom panels) vs. concentration for a positive compound with 18 or 27 subjects for each of the scenario simulated - typical FTIH design, modified design with moxifloxacin PD measurements and modified design with moxifloxacin PK and PD measurements. The dotted vertical line in each of the graphs represent the concentration at which the hypothetical drug effects were confined to. The two dotted horizontal lines in the top panels indicate the positive drug effect of 10 msec QT prolongation and negative effect of 2 msec increase. Each plot also contains the median, 5th and 95th percentile of the predicted QT prolongation as well as the probability of 10 msec increase.

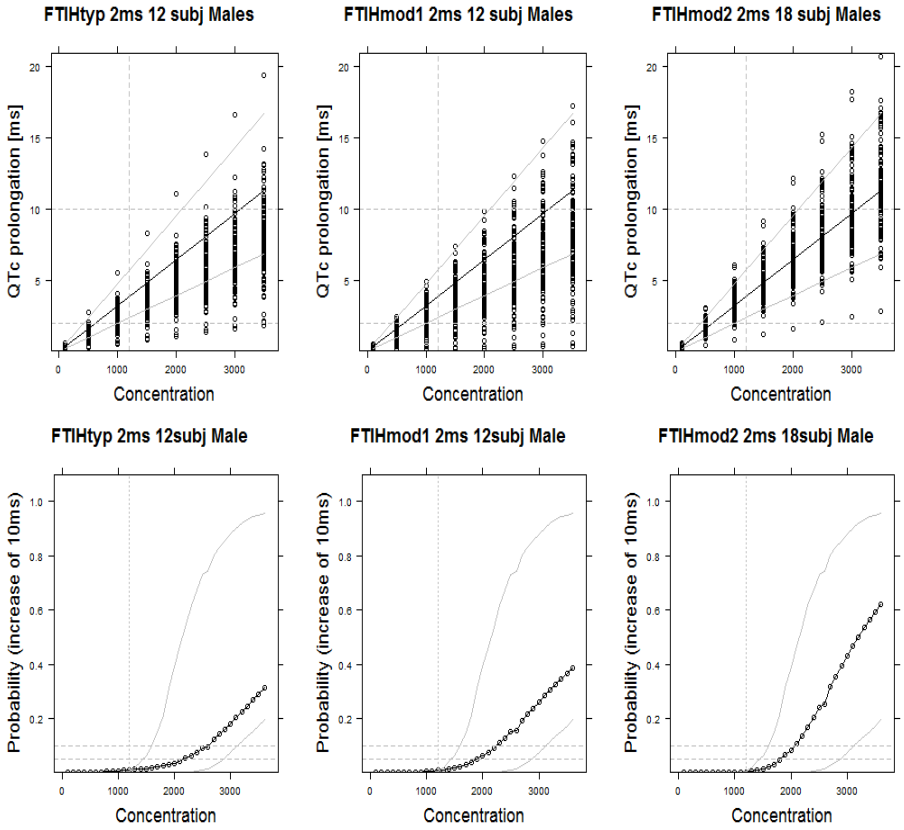


Figure 4. Relationship between QT prolongation (top panels) and probability of ≥10 msec increase (bottom panels) vs. concentration for a negative compound with 12 or 18 subjects for each of the scenario simulated - typical FTIH design, modified design with moxifloxacin PD measurements and modified design with moxifloxacin PK and PD measurements. The dotted vertical line in each of the graphs represent the concentration at which the hypothetical drug effects were confined to. The two dotted horizontal lines in the top panels indicate the positive drug effect of 10 msec QT prolongation and negative effect of 2 msec increase. Each plot also contains the median, 5th and 95th percentile of the predicted QT prolongation as well as the probability of 10 msec increase.

Type I and type II error rates

Finally, Tables 3a and b displays the specificity and sensitivity rates that are compared between the double delta method used in a traditional TQT and Bayesian approach suggested for the FTIH studies using the same sample sizes, namely 12, 18 and 27. The trends are plotted in Figure 5.

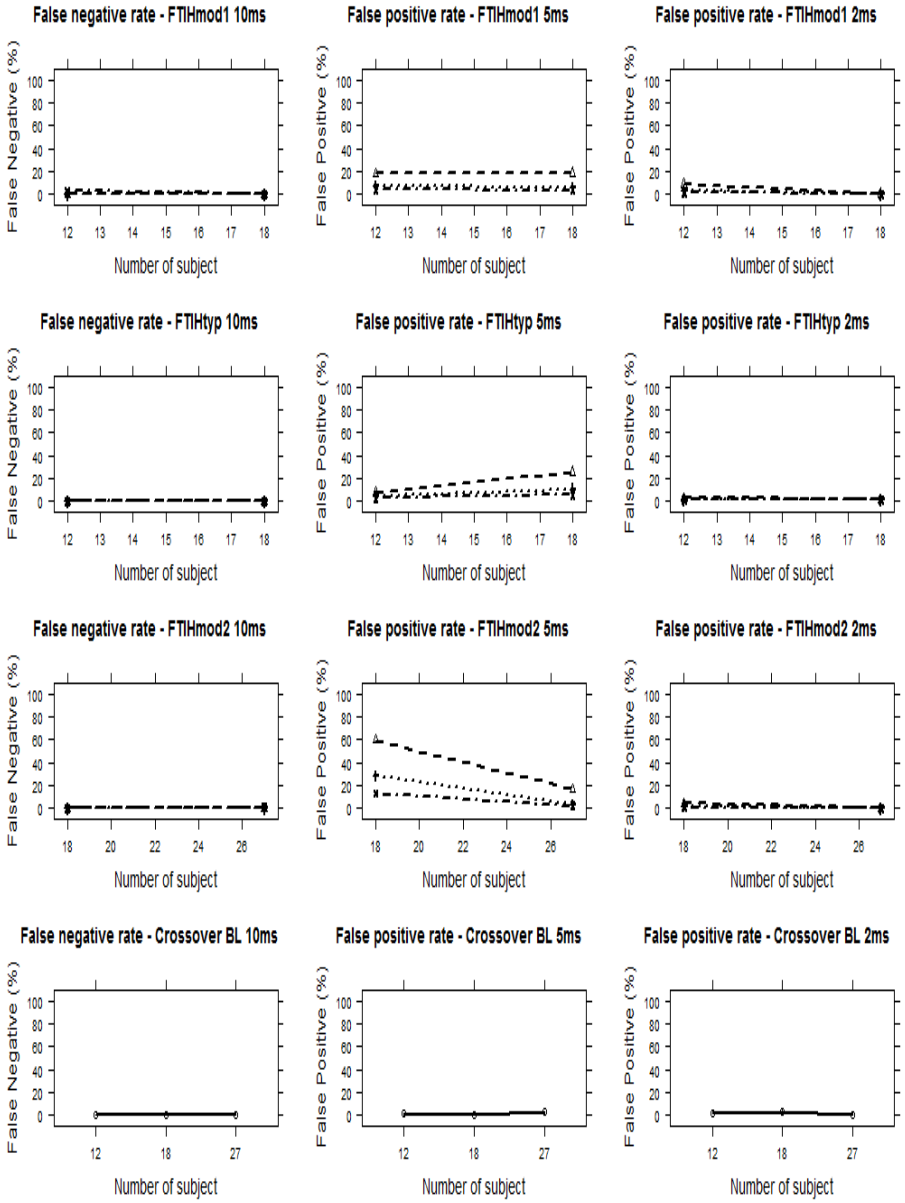


Figure 5. False negative and positive rates for the Bayesian approach used in the FTIH studies with various study scenarios - typical FTIH design (top row), modified 1 design (second row), modified 2 design (third row) and number of subjects. Δ, + and x dash lines describe the results using different definitions of threshold of detection at 99%, 95% and 90% respectively. The last row contains false positive and negative rates using double delta method with the same number of subjects.

4 msec variability on SLP		CRbl 12	CRbl 18	CRbl 27
DD	*Specificity*	0.96	0.915	0.96
	Sensitivity	1	1	1

4 msec variability of SLP		Typ (12)	Typ (18)	Mod 1 (12)	Mod 1 (18)	Mod 2 (18)	Mod 2 (27)
BUGS (P >= 99%)	*Specificity*	0.95	0.87	0.87	0.91	0.68	0.92
	Sensitivity	1	1	1	1	1	1
BUGS (P >= 95%)	*Specificity*	0.97	0.94	0.95	0.97	0.85	0.98
	Sensitivity	1	1	1	1	1	1
BUGS (P >= 90%)	*Specificity*	0.98	0.97	0.97	0.98	0.93	0.99
	Sensitivity	1	1	0.97	1	1	0.99

Table 3. Specificity and sensitivity tabulation results the FTIH simulations with 4 msec variability applied to the slope parameter for the double-delta method (top) and the Bayesian approach (bottom). Results are shown for detection thresholds at 99%, 95% and 90%.

11.4 DISCUSSIONS

Our study has demonstrated that FTIH studies can be used to assess the probability of drug-induced QTc prolongation. In contrast to current belief regarding the requirement for statistical power of a study to detect 5 msec of prolongation, the false-positive rates in the newly proposed FTIH studies are in general lower than what is observed in a typical TQT study. Moreover, the current execution of TQT studies does not necessarily guarantee accuracy and precision of the effect size which does not satisfy the intention of conducting a stand-alone study dedicated to detect any proarrhythmic property of new compounds. An advantage to using the Bayesian PKPD approach to analyse the proarrhythmic property is its ability to describe the prolongation effect over the therapeutic range or specific dose levels. Since FTIH trial is a mandatory step in the drug development process, to increase the number of ECG measurements taken to provide a larger PD sample size will provide an initial assessment of the compound's safety profile. This can in turn support the strategic decision to proceed or not with development.

In our study, we did not find a difference in the outcomes with and without the inclusion of a positive-control arm. Given that benchmark moxifloxacin data may not explicitly reflect individual sensitivity to the effects of a NCE in

this instance, a comprehensive assessment of the propensity for QTc-interval prolongation may be achieved by further integration with preclinical findings, from *in vitro* experiments and *in vivo* studies. This further call for the need to generate safety data in an integrated manner, in which drug exposure and potential biomarkers are analysed using a model-based approach to ensure the results are translatable and comparable. However, there is still no consensus among regulatory authorities whether preclinical results can predict the risk of QTc prolongation in humans and guide the early clinical QTc assessment strategy [9, 16-18], as evident from the ICH S7A and S7B guidance [19, 20].

The Bayesian approach is limited to improving the analytical aspect to detect QTc-interval prolongation of a compound and thus 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 6 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.

In general, most forms of ECG measurements (manual, automated, computer-assisted (CA)) are able to demonstrate assay sensitivity of moxifloxacin in healthy volunteers. However the measurement variability differs greatly with manual and CA measurements at 5-28% greater than that associated with automated methods [21]. 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 [22, 23], identifying sources of variability [24], new algorithms [25-30] as well as the issues of reproducibility [31-33].

In 2001, the US Food and Drug Administration (FDA) even requires digital ECGs from TQT be submitted into its Warehouse and encourages other recordings as part of new drug application (NDA) dossiers for review and approval purposes [34]. 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

[34], determine the underlying physiological variation, and use as prior distributions for future Bayesian PKPD analysis for QTc-interval prolongation.

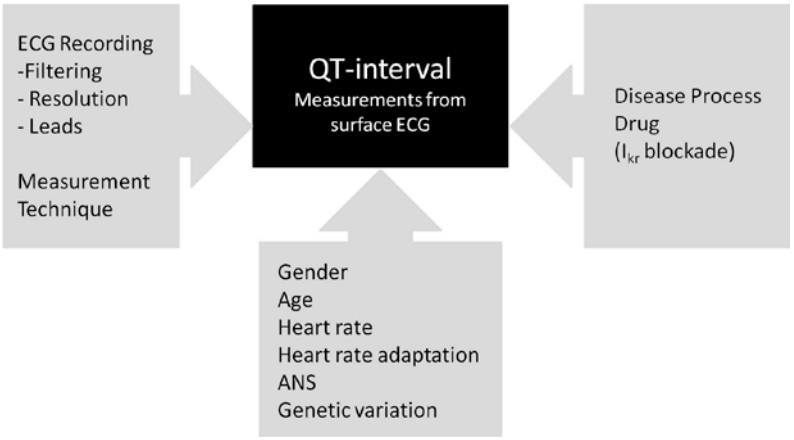


Figure 4. Summary of the potential areas of concern where significant impact can be made to the observed QT-interval data, including the algorithm used for ECG recordings, the technique selected to correct measurement bias, the nature of the disease that influence ECG values, and various other covariates that can impact the recordings [35].

There are occasions when traditional TQT studies cannot be conducted in special populations such as testing anti-cancer therapies in patients [36]. Due to the fact that our mechanistic PKPD approach can detect prolongation at any observable concentration range, and does not necessarily require the inclusion of a positive control arm, it can be applied to any trial populations where only therapeutic doses can be applied. Our results showed that evidence for the drug concentration-QTc prolongation relationship can be accurately derived from FTIH studies, which provides safety information regarding the compound early in the clinical development process. In the Bayesian analysis approach, biological and mechanism-based elements are differentiated thus making the model translatable. Furthermore, we have demonstrated that PKPD relationships can eliminate statistical flaws associated with the double-delta method, but most importantly decrease type I and II error rates in the traditional approach. Thus by extension, the results clearly challenged the need and relevance of a TQT study.

REFERENCES

[1] Chain AS, Krudys KM, Danhof M, Della Pasqua O. Assessing the Probability of Drug-Induced QTc-Interval Prolongation During Clinical Drug Development. Clin Pharmacol Ther 2011 Dec;90(6):867-75.

[2] Guidance for Industry: E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrythmic Potential for Non-Antiarrythmic Drugs. 2005.

[3] Morganroth J. A definitive or thorough phase 1 QT ECG trial as a requirement for drug safety assessment. J Electrocardiol 2004 Jan;37(1):25-9.

[4] Rohatagi S, Carrothers TJ, Kuwabara-Wagg J, Khariton T. Is a thorough QTc study necessary? The role of modeling and simulation in evaluating the QTc prolongation potential of drugs. J Clin Pharmacol 2009 Nov;49(11):1284-96.

[5] Sethuraman V, Sun Q. Impact of baseline ECG collection on the planning, analysis and interpretation of 'thorough' QT trials. Pharm Stat 2009 Apr-Jun;8(2):113-24.

[6] Tsong Y, Zhong J. Multiple comparisons of repeated measured response: issues of assessment of prolongation of QT interval in thorough QT trials. J Biopharm Stat 2010 May;20(3):604-14.

[7] Bouvy JC, Koopmanschap MA, Shah RR, Schellekens H. The Cost-Effectiveness of Drug Regulation: The Example of Thorough QT/QTc Studies. Clin Pharmacol Ther 2012 Feb;91(2):281-8.

[8] Malik M, Garnett CE, Zhang J. Thorough QT Studies: Questions and Quandaries. Drug Saf 2010 Jan 1;33(1):1-14.

[9] Cavero I, Crumb W. The use of electrocardiograms in clinical trials: a public discussion of the proposed ICH E14 regulatory guidance. April 11-12, 2005, Bethesda, MD, USA. Expert Opin Drug Saf 2005 Jul;4(4):795-9.

[10] Taubel J, Wong AH, Naseem A, Ferber G, Camm AJ. Shortening of the QT Interval After Food Can Be Used to Demonstrate Assay Sensitivity in Thorough QT Studies. J Clin Pharmacol 2011 Nov 8.

[11] Florian JA, Tornoe CW, Brundage R, Parekh A, Garnett CE. Population pharmacokinetic and concentration--QTc models for moxifloxacin: pooled analysis of 20 thorough QT studies. J Clin Pharmacol 2011 Aug;51(8):1152-62.

[12] Beal S, Sheiner, L. NONMEM user's guide NONMEM project group. San Francisco; 1999.

[13] Lunn DJ, Thomas, A., Best, N., Spiegelhalter, D. WinBUGS - A Bayesian modelling framework: concepts, structure, and extensibility. Stat Comput 2000;10:325-37.

[14] Spiegelhalter D, Thomas, A., Best, N. & Gilks W. BUGS: Bayesian Inference Using Gibbs Sampling, Version 1.4.1 Cambridge, UK: MRC Biostatistics Unit.

[15] R Development Core Team, R: a language and environment for statistical compting. In: computing Rffs, editor. Vienna, Austria; 2010.

[16] Cavero I, Crumb W. ICH S7B draft guideline on the non-clinical strategy for testing delayed cardiac repolarisation risk of drugs: a critical analysis. Expert Opin Drug Saf 2005 May;4(3):509-30.

[17] Shah RR. Drugs, QTc interval prolongation and final ICH E14 guideline : an important milestone with challenges ahead. Drug Saf 2005;28(11):1009-28.

[18] Darpo B, Nebout T, Sager PT. Clinical evaluation of QT/QTc prolongation and proarrhythmic potential for nonantiarrhythmic drugs: the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use E14 guideline. J Clin Pharmacol 2006 May;46(5):498-507.

[19] Guidance for Industry: Safety pharmacology studies for human pharmaceuticals. 2001.

[20] Guidance for Industry: The nonclinical evaluation of the potential for delayed ventricular repolarization (QT interval prolongation) by human pharmaceuticals. 2004.

[21] Fosser C, Duczynski G, Agin M, Wicker P, Darpo B. Comparison of manual and automated measurements of the QT interval in healthy volunteers: an analysis of five thorough QT studies. Clin Pharmacol Ther 2009 Nov;86(5):503-6.

[22] Zhang J. Testing for positive control activity in a thorough QTc study. J Biopharm Stat 2008;18(3):517-28.

[23] Chiang AY, Bass AS, Cooper MM, Engwall MJ, Menton RG, Thomas K. ILSI-HESI cardiovascular safety subcommittee dataset: an analysis of the statistical properties of QT interval and rate-corrected QT interval (QTc). J Pharmacol Toxicol Methods 2007 Sep-Oct;56(2):95-102.

[24] Agin MA, Aronstein WS, Ferber G, Geraldes MC, Locke C, Sager P. QT/QTc prolongation in placebo-treated subjects: a PhRMA collaborative data analysis. J Biopharm Stat 2008;18(3):408-26.

[25] Sarapa N, Gussak I, Vajdic B, George S, Hadzievski L, Francom SF, et al. Comparison of QTinno, a fully automated electrocardiographic analysis program, to semiautomated electrocardiographic analysis methods in a drug safety study in healthy subjects. J Electrocardiol 2009 Jul-Aug;42(4):358-66.

[26] Khawaja A, Petrovic R, Safer A, Baas T, Dossel O, Fischer R. Analyzing Thorough QT Study 1 & 2 in the Telemetric and Holter ECG Warehouse (THEW) using

Hannover ECG System HES : A validation study. Comput Cardiol (2010) 2010;37:349-52.

[27] Martinez JP, Laguna P, Olmos S, Pahlm O, Petersson J, Sornmo L. Accuracy of QT measurement in the EASI-derived 12-lead ECG. Conf Proc IEEE Eng Med Biol Soc 2006;1:3986-9.

[28] Martinez JP, Laguna P, Olmos S, Pahlm O, Pettersson J, Sornmo L. Assessment of QT-measurement accuracy using the 12-lead electrocardiogram derived from EASI leads. J Electrocardiol 2007 Apr;40(2):172-9.

[29] Hnatkova K, Smetana P, Toman O, Bauer A, Schmidt G, Malik M. Systematic comparisons of electrocardiographic morphology increase the precision of QT interval measurement. Pacing Clin Electrophysiol 2009 Jan;32(1):119-30.

[30] Salvi V, Karnad DR, Panicker GK, Natekar M, Hingorani P, Kerkar V, et al. Comparison of 5 methods of QT interval measurements on electrocardiograms from a thorough QT/QTc study: effect on assay sensitivity and categorical outliers. J Electrocardiol 2011 Mar-Apr;44(2):96-104.

[31] Darpo B, Fossa AA, Couderc JP, Zhou M, Schreyer A, Ticktin M, et al. Improving the precision of QT measurements. Cardiol J 2011;18(4):401-10.

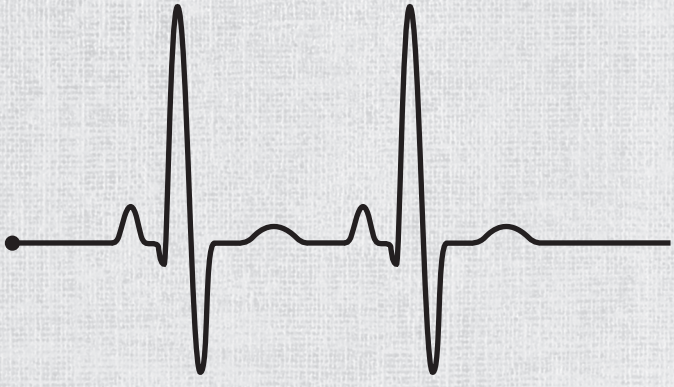
[32] Couderc JP, Garnett C, Li M, Handzel R, McNitt S, Xia X, et al. Highly automated QT measurement techniques in 7 thorough QT studies implemented under ICH E14 guidelines. Ann Noninvasive Electrocardiol 2011 Jan;16(1):13-24.

[33] Postema PG, De Jong JS, Van der Bilt IA, Wilde AA. Accurate electrocardiographic assessment of the QT interval: teach the tangent. Heart Rhythm 2008 Jul;5(7):1015-8.

[34] Sarapa N. Quality assessment of digital annotated ECG data from clinical trials by the FDA ECG Warehouse. Expert Opin Drug Saf 2007 Sep;6(5):595-607.

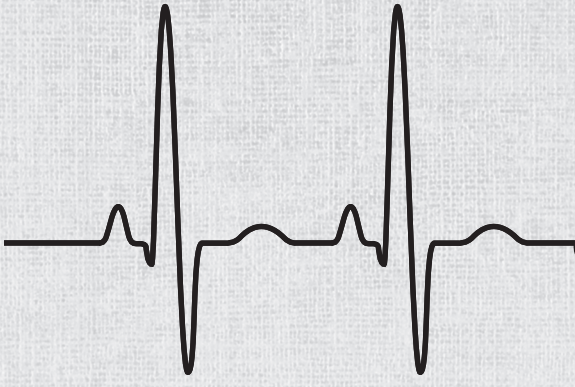
[35] Totterman M. iCardiac Technologies Inc, The Telemetric and Holter Warehouse Project, January,2008.

[36] Rock EP, Finkle J, Fingert HJ, Booth BP, Garnett CE, Grant S, et al. Assessing proarrhythmic potential of drugs when optimal studies are infeasible. Am Heart J 2009 May;157(5):827-36.



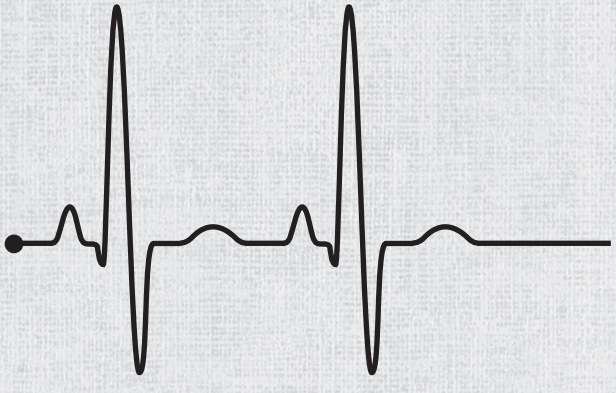
"Only those who will risk going too far can possibly find out how far one can go."

T.S. Elliot



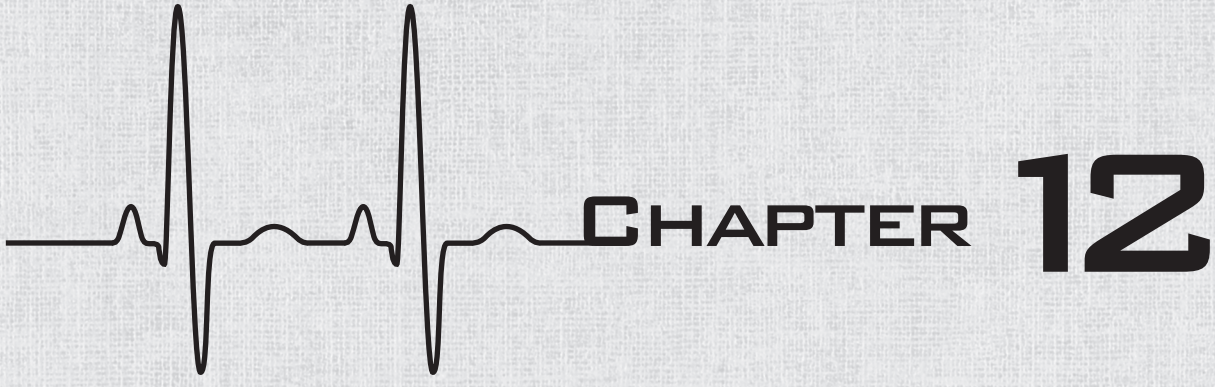
SECTION V

CONCLUSIONS & PERSPECTIVE



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 *in vitro* and *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 **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.

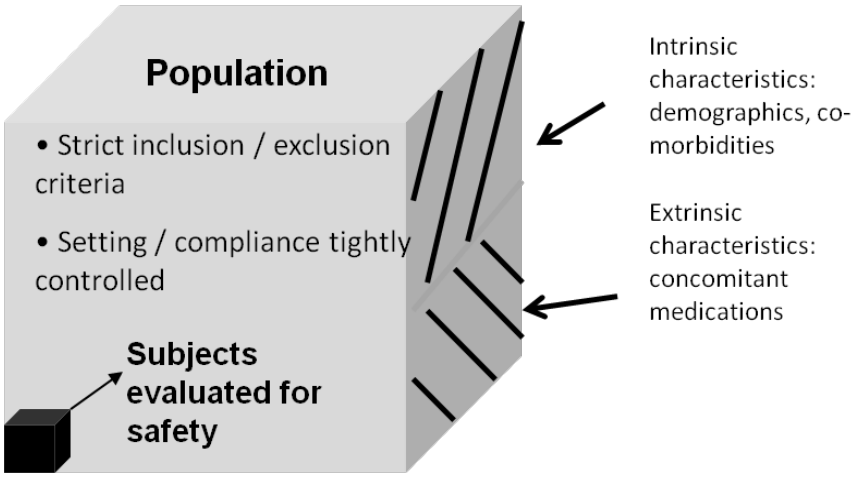


Figure 1: Clinical trials and in particular the TQT trial design yields estimates of drug effects on a restricted subset of the population. The expectation that such a study will generate accurate evidence of the true safety profile of a new chemical entity in the overall population (with different demographic 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).

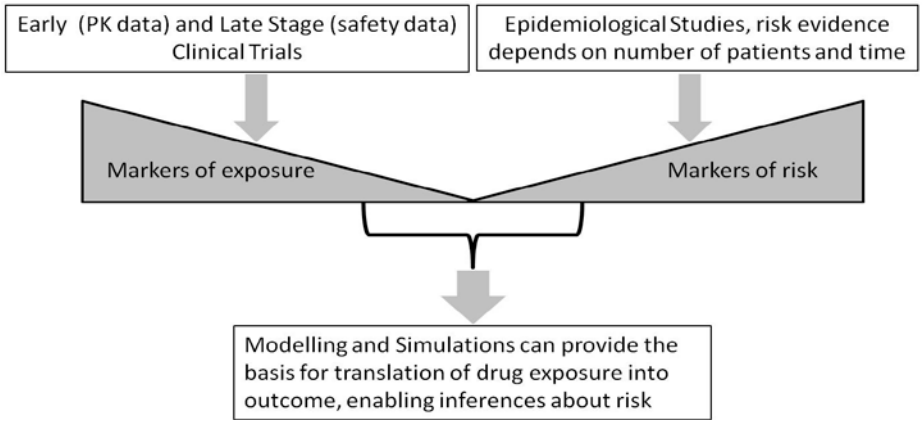


Figure 2. Inferential methods are essential to establish causality during signal detection. The inclusion of concentration-effect relationships into pharmaco-epidemiological research can provide more accurate, quantitative estimates of drug-induced effects. Different experimental and non-experimental 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.

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 $\geq$ 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

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 C*hapters 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 C*hapter 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 *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 *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

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 C*hapter 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 C*hapter 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 $\geq$ 10msec 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 C*hapter 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 C*hapters 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 C*hapter 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.

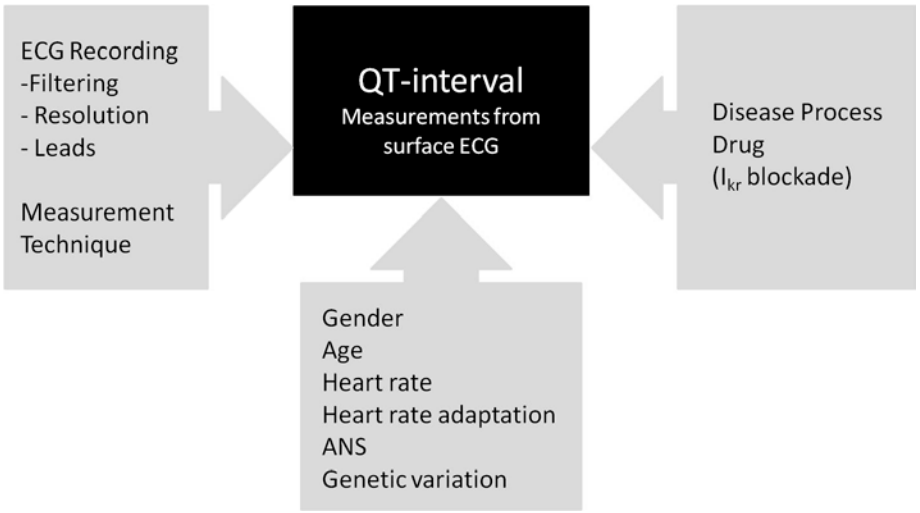


Figure 3. Various factors that affect QT-interval measurements. A combination of contributors should be taken into considerations when assessing QT/QTc-interval prolongation including intrinsic and extrinsic characteristics [22].

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 issued 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

[1] Kowey PR, Malik, M. The QT interval as it relates to the safety of non-cardiac drugs. European Heart Journal Supplements 2007;9 (Supplement G).

[2] Letsas KP, Letsas, GP. Drug-Induced Proarrhythmia: QT Interval Prolongation and Torsade de Pointes. Hospital Chronicles 2011;6(3):118-22.

[3] Piccini JP, Whellan DJ, Berridge BR, Finkle JK, Pettit SD, Stockbridge N, et al. Current challenges in the evaluation of cardiac safety during drug development: translational medicine meets the Critical Path Initiative. Am Heart J 2009 Sep;158(3):317-26.

[4] Turner JR. Drug safety, medication safety, patient safety: An overview of recent FDA guidances and initiatives. Regulatory Rapporteur 2009;6(4).

[5] Pollard CE, Valentin JP, Hammond TG. Strategies to reduce the risk of drug-induced QT interval prolongation: a pharmaceutical company perspective. Br J Pharmacol 2008 Aug;154(7):1538-43.

[6] Anon. ICH E14 The clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs. 2005.

[7] Anon. ICH Topic S7A Safety Pharmacolog Studies for Human Pharmaceuticals. 2000.

[8] Anon. ICH Topic S7B The nonclinical Evaluation of the Potential for delayed Ventricular Repolarization (QT Interval Prolongation) by Human Pharmaceuticals. 2005.

[9] Pater C. Methodological considerations in the design of trials for safety assessment of new drugs and chemical entities. Curr Control Trials Cardiovasc Med 2005 Feb 3;6(1):1.

[10] Turner JR. Analyzing Safety Data. Key Statistical Concepts in Clinical Trials for Pharma: Springer; 2012. p. 11-21.

[11] Pepe M, Longton G, Janes H. Estimation and Comparison of Receiver Operating Characteristic Curves. Stata J 2009 Mar 1;9(1):1.

[12] Cavero I. Using pharmacokinetic/pharmacodynamic modelling in safety pharmacology to better define safety margins: a regional workshop of the Safety Pharmacology Society. Expert Opin Drug Saf 2007 Jul;6(4):465-71.

[13] Fleury A, Lave T, Jonsson F, Schmitt M, Hirkaler G, Polonchuk L, et al. A pharmacokinetic-pharmacodynamic model for cardiovascular safety assessment of R1551. J Pharmacol Toxicol Methods 2010 Jan-Feb;63(1):123-33.

[14] Langdon G, Davis JD, McFadyen LM, Dewhurst M, Brunton NS, Rawal JK, et al. Translational pharmacokinetic-pharmacodynamic modelling; application to cardiovascular safety data for PF-00821385, a novel HIV agent. Br J Clin Pharmacol 2010 Apr;69(4):336-45.

[15] Ollerstam A, Visser SA, Duker G, Forsberg T, Persson AH, Nilsson LB, et al. Comparison of the QT interval response during sinus and paced rhythm in conscious and anesthetized beagle dogs. J Pharmacol Toxicol Methods 2007 Sep-Oct;56(2):131-44.

[16] Bonate PL, Russell T. Assessment of QTc prolongation for non-cardiac-related drugs from a drug development perspective. J Clin Pharmacol 1999 Apr;39(4):349-58.

[17] Black N. Why we need observational studies to evaluate the effectiveness of health care. BMJ 1996 May 11;312(7040):1215-8.

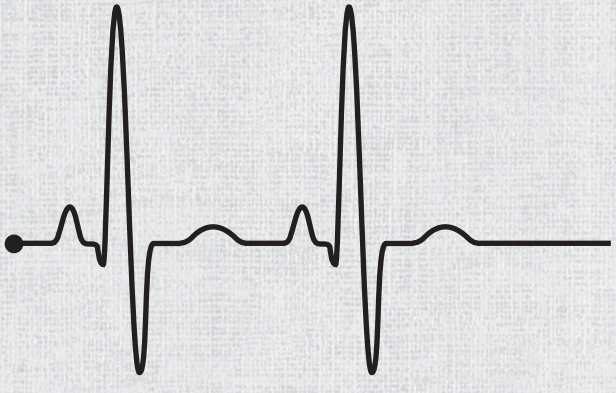
[18] Baker SG, Cook NR, Vickers A, Kramer BS. Using relative utility curves to evaluate risk prediction. J R Stat Soc Ser A Stat Soc 2009 Oct 1;172(4):729-48.

[19] Vickers AJ, Cronin AM, Kattan MW, Gonen M, Scardino PT, Milowsky MI, et al. Clinical benefits of a multivariate prediction model for bladder cancer: a decision analytic approach. Cancer 2009 Dec 1;115(23):5460-9.

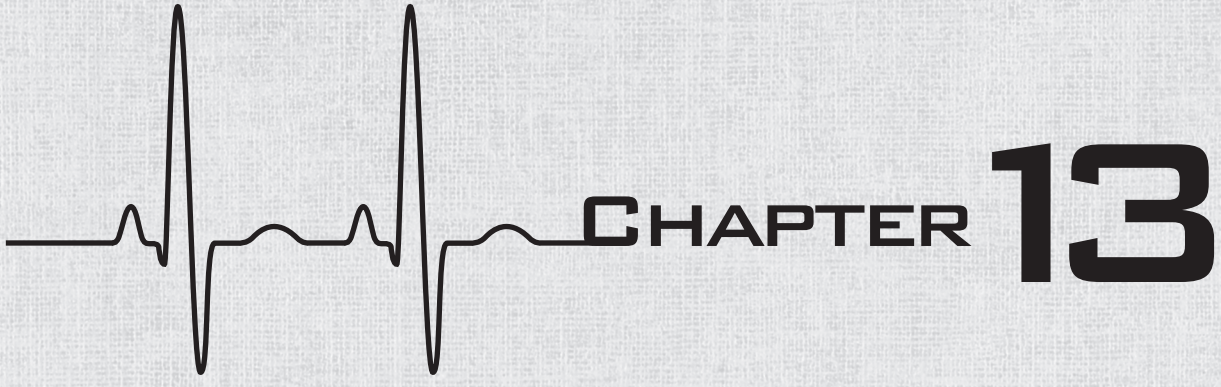
[20] CIOMS. Benefit-Risk Balance for Marketed Drugs: Evaluating Safety Signals. Geneva: CIOMS Working Group V; 1998.

[21] Walker S, Philips, L., Cone, M. Benefit-Risk Assessment Model for Medicines: Developing a Structured Approach to Decision Making. Washington DC: CMR International Institute for Regulatory Science; 2005.

[22] Totterman M. iCardiac Technologies Inc, The Telemetric and Holter Warehouse Project, January,2008.



Nederlandse Samenvatting
(Synopsis in Dutch)



Chapter 13

Mind The Gap: De voorspelling van cardiovasculaire veiligheid tijdens de ontwikkeling van nieuwe geneesmiddelen – Samenvatting, conclusies en vooruitblik

13.1 HET REGULATOIR KADER EN DE VERTALING VAN DE PRO-ARITMISCHE ACTIVITEIT VAN GENEESMIDDELEN NAAR RISICO

Verscheidene intrinsieke en extrinsieke factoren, waaronder genetische predispositie, kunnen het zogenoemde QT (of QTc)-interval in een electrocardiogram verlengen. Hoewel QT verlenging niet gevaarlijk hoeft te zijn, wordt het in verband gebracht met hartritmestoornissen en plotselinge hartdood [1, 2]. Aanwijzingen dat een geneesmiddel het QTc interval kan verlegen is in het algemeen een reden tot zorg voor regelgevende autoriteiten [3, 4]. Ondanks inspanningen om het effect van geneesmiddelen op de hartgeleiding in een vroeg stadium van ontwikkeling te identificeren, is QTc verlenging de op-één-na meest voorkomende reden om een geneesmiddel uit de handel te nemen. Verder bljkt het aantal geneesmiddelen die beschouwd worden een risicovol QT profiel te tonen geleidelijk toe te nemen.

Deze recente "*farmacoepidemie*" in ogenschouw nemend, is het van belang het QT-verlengingspotentiaal te voorspellen en de klinische relevantie ervan zo vroeg mogelijk vast te stellen, zodat wijgering of vertraging van de geneesmiddelregistratie en beperkingen in het gebruik van het geneesmiddel vermeden kunnen worden [5].

Naast de ontwikkeling van onderzoeksprotocollen voor het vaststellen van cardiovasculaire veiligheid in dierproeven worden in het klinisch onderzoek de zogenoemde E14 richtlijnen toegepast, welke zijn opgesteld door de International Conference on Harmonisation (ICH) [6]. Gelijksoortige richtlijnen (nummers S7A en S7B) zijn ook ontwikkeld voor onderzoek in dieren [7, 8]. De E14 richtlijnen bevelen een aparte studie aan die de cardiovasculaire veiligheid moet aantonen. Deze zogenoemde grondige QT studie (thorough QT, TQT) is zodanig ontworpen dat een eventuele QT verlenging van 5 milliseconden aangetoond kan worden. Indien het

bijbehorende éénzijdige betrouwbaarheidsinterval de 10 milliseconden niet overschrijdt, wordt dit geïnterpreteerd als een gebrek aan bewijs voor een risicovolle QT verlenging. Alhoewel de aanbevolen invoering van risico-evaluatie goed bedoeld is, zijn de aanbevelingen allesbehalve perfect.

Dit proefschrift belicht drie belangrijke tekortkomingen in de huidige aanpak, welke de evaluatie en ontwikkeling van geneesmiddelen ondermijnen en daardoor de vermindering van de bestaande medische behoefte belemmeren. De eerste tekortkoming betreft **de aannames die ten grondslag liggen aan de vertaling van farmacologische effecten in experimenten en klinische studies naar die in de medische praktijk**. De uitwerkingen van experimentele factoren en processen wordt over het hoofd gezien, wat leidt tot een onvolledig begrip van de betekenis van farmacologische eigenschappen en het vertaling ervan naar kwantificeerbare risico's. De tweede tekortkoming heeft te maken met **het verschil tussen associatie en causaliteit en de biologische implicaties hiervan**. De keuze voor het QT interval als maat voor het risico op hartritmestoornissen zonder hierbij de relatie tussen geneesmiddelblootstelling en farmacologisch effect te betrekken bemoeilijkt het vaststellen van een causaal verband tussen eerstgenoemden. Eenvoudige principes van signaaldetectie, inclusief discrimatiedrempel zoals aangegeven door ROC curves en begripsvaliditeit worden veronachtzaamd in onderzoeksprotocollen. De derde tekortkoming is **het ontbreken van een methode om afzonderlijke aanwijzingen van de risico's en baten van de behandeling samen te smeden** tot één waarde ofwel "evidence synthesis". De nadruk op het genereren van aanwijzingen zonder deze vervolgens te integreren leidt tot suboptimale protocollen en onjuiste beslissingscriteria. Dit heeft tot gevolg dat de ontwikkeling van een geneesmiddel onterecht beëindigd kan worden. De huidige methodes staan een integratieve benadering zelfs niet toe: noch van het te verwachten risico, noch van de effectiviteit van maatregelen om het risico te verminderen.

Idealiter zouden experimenten in celpreparaten reeds een goede voorspelling verschaffen van de te verwachten klinische effecten, zowel vanuit het kwalitatieve perspectief (de voorspelling luidt: ofwel QT interval verlenging vindt plaats ofwel het vindt niet plaats) als vanuit het kwantitatieve (een voorspelling van de QT verlenging per eenheid verandering van de geneesmiddelconcentratie). Verschillende auteurs hebben methodes voorgesteld om een kwalitatieve voorspellende relatie te

identificeren [9, 10]. Deze relaties zijn echter gebaseerd op een empirisch gemaakt onderscheid tussen diverse bevindingen. Het probleem hierbij is dat de classificatie van geneesmiddelen met een klein QT verlengingseffect moeilijk is. Verder ontbreken er kwantitatieve benaderingen, die de integratie nastreven van de resultaten die verkregen worden in de diverse stadia van geneesmiddelontwikkeling.

Hoewel het werk in dit proefschrift de voorspelling van QT effecten met gebruikmaking van *in vivo* (uit levende organismen) verkregen informatie behelst, blijft ook de ontwikkeling van voorspellende methoden met behulp van *in vitro* (c.q. afkomstig uit experimenten met celpreparaten) gegevens van groot belang. De mogelijkheid tot voorspelling met op basis van *in vitro* verkregen maatstaven zou het vervolgproces van experimenten en beslissingen al in een vroeg stadium kunnen vereenvoudigen. Dit betekent een mogelijke beperking van het aantal afgewezen kandidaat-moleculen.

De nadruk ligt in dit werk echter op het *in vivo* stadium van geneesmiddelontwikkeling. Barrières voor de vertaling van onderzoeksgegevens verkregen in dit stadium naar de late klinische en praktijksituaties zijn overbrugbaar, zeker wanneer het gegevens uit vroeg-klinische studies betreft: hier immers worden de geneesmiddelblootstelling en cardiovasculaire signalen altijd uitvoerig gemeten. De belangrijkste doelstelling van dit proefschrift is daarom de identificatie van ontbrekende factoren in de genoemde vertaalslag. Deze factoren zijn het gevolg van verschillen in de gecontroleerde omstandigheden waarin studies worden uitgevoerd, met name de verschillen in de toegepaste inclusie- en exclusiecriteria. Aan dit werk ligt de vaststelling ten grondlag dat aan de medische behandeling met geneesmiddelen altijd risico's verbonden zijn. Beslissingen aangaande de geschiktheid van een molecuul voor verdere ontwikkeling tot geneesmiddel alsmede het therapeutisch gebruik van het ontwikkelde geneesmiddel moeten daarom gebaseerd zijn op een kwantitatieve bepaling van het risico in de patientendoelgroep. In plaats van te vertrouwen op de veronderstelde algemene geldigheid en extrapoleerbaarheid van individuele onderzoeksprotocollen, is het beter om uit te gaan van risicoanalyses waarin de verschillen tussen experiment en praktijk in acht genomen worden.

Een belangrijke bevinding die naar voren komt uit dit werk is dat het samenvoegen van bestaande resultaten doeltreffender kan zijn dan het

genereren van nieuwe gegevens. Wanneer deze synthese plaatsvindt met de juiste klinische en wetenschappelijke intentie en met toepassing van geschikte statistische methoden, kan zij een betere basis vormen voor besluitvorming. Tegelijkertijd respecteert zij praktische en ethische bezwaren die verbonden zijn aan het onderzoek van geneesmiddelveiligheid in de mens. Dit is speciaal van belang wanneer het gaat om veiligheidssignalen die met een lage frequentie voorkomen of wanneer het risiko gedurende of na chronische behandeling ontstaat.

De genoemde tekortkomingen, de hoge kosten verbonden aan het uitvoeren van een TQT studie en de noodzaak om de klinische betekenis van vroege veiligheidssignalen op een nauwkeurige mannier in te kunnen schatten vormen tezamen een pleidooi voor het beproeven van nieuwe strategieen voor de evaluatie van de cardiovasculaire veiligheid van geneesmiddelen. Bovendien is het essentieel dat de regelgeving op dit gebied een afspiegeling wordt van de wetenschappelijke kennis. Dit is evident niet het geval voor de ICH E14 richtlijnen, welke nog gestoeld zijn op de wetenschappelijke en regelgevingsgeest zoals die meer dan tien jaar geleden heerste. In die tijd waren de kwantitatieve farmacologie en farmacometrie als klinische disciplines nog in ontwikkeling. Nieuwe, robuuste methodes zouden in overweging moeten worden genomen en in de regelgeving opgenomen. Richtlijnen en regelgevingsbeleid zouden voortschrijdende inzichten moeten belichamen. Inferentiemethoden gebaseerd op modellen en simulaties blijven in veiligheidsbepalingen nog ondergebruikt. De grote hoeveelheid data die gegenereerd wordt tijdens de klinische ontwikkelingsfase en de beschikbaarheid van gegevens over geneesmiddelmoleculen met vergelijkbare farmacologische eigenschappen worden niet systematisch – vanaf de vroege klinische studies tot aan de veiligheidsbewakking na toelating – ingezet.

De in dit proefschrift voorgestelde nieuwe denkwijze stelt *"evidence synthesis"* centraal en beschouwt het genereren van nieuwe informatie als een bevestigingsstap in het proces dat leidt van hypothese naar kennis. Statistische methodes vormen de gereedschappen voor deze synthese, die in meer specifieke zin de vertaling van veiligheidssignalen naar risico's behelst. Het hier beschreven onderzoek introduceert bovendien farmacologische concepten in een epidemiologische context en verkleint daarmee het onderscheid tussen geneesmiddelontwikkeling en gezondheidszorg.

Een direct gevolg van de hier gekozen benadering is een accentverschuiving in de analyse van veiligheidssignalen, van associatie naar causaliteit.

Gestoeld op bovenstaande principes geeft dit proefschrift een aanzet tot 1) het terugdringen van het aantal afgewezen kandidaat-moleculen gedurende het screeningsonderzoek en 2) de ontwikkeling van betere onderzoeksprotocollen voor het vaststellen de het pro-arrhythmische effect van nieuwe geneesmiddelen. **Hoofdstuk 1** verschaft inzicht in het huidige proces van de selectie en ontwikkeling van kandidaat-moleculen. De voorspellende waarde van onderzoeksprotocollen wordt aan een kritische blik onderworpen en de voorwaarden voor de vertaling van bevindingen worden beschouwd in relatie tot de experimentele procedures en data-analysemethodes waaruit zij voortkomen. Bij de data analyse wordt speciaal gelet op de keuze van parameterisering. Deze wordt nader onderzocht in **hoofdstuk 2**. Het gebruik van zogenaamde hierarchische modellen vormt een geschikte basis voor de vertaling van geneesmiddeleffecten naar praktische risico's. In dat opzicht onderstreept dit werk niet alleen het belang van alternatieve methodologie voor het kwantificeren van QT (of QTc) intervalverlenging, maar beschrijft het ook de eisen voor de parameterisatie van de hiervoor gebruikte modellen. Vervolgens werd een dergelijk model dat het verband tussen geneesmiddelconcentratie en effect beschrijft geëvalueerd. In de daaropvolgende stap lag de nadruk op de waarde van een strategie voor de overbrugging van de resultaten uit clinische studies naar de doel patiëntenpopulatie in de medische praktijk. Deze strategie kreeg de benaming Not-in-Trial Simulaties (NITS). NITS verdient aanbeveling als methode voor besluitvorming en risicomanagement. Tenslotte werd de waarde onderzocht van enkele aanpassingen in de standaard onderzoeksprotocollen voor de screening van kandidaat-moleculen. Ons onderzoek ging ook in op de specificiteit en sensitiviteit van een standaard gehanteerde drempelwaarde voor QT verlenging (nl. 10 milliseconden). Voor waarden boven deze drempel wordt aangenomen dat er een klinisch relevant cardiovasculair risico bestaat. In principe kan men ervan uit gaan, dat de voorspellende waarde van resultaten uit vroege ontwikkelingsfases vastgesteld kunnen worden uit de zogenaamde receiver operating characteristic (ROC) curves [11]. Hoewel een formele ROC analyse buiten beschouwing bleef, werden enkele tekortkomingen in de aannames van gerelateerd beleid en richtlijnen aangetoond.

13.2 BAYESIAANSE HIERARCHISCHE MODELLEN

In **hoofdstuk 3** werd getoond hoe met een Bayesiaans hierarchisch model het verband tussen geneesmiddelconcentratie en de verlenging van het QTc interval beschreven kan worden. Deze toepassing werd gedemonstreerd aan de hand van drie referentiegeneesmiddelen *d,l*-sotalol, moxifloxacine en grepafloxacine. Het gebruik van een geïntegreerde modelmatige benadering stelde ons in staat het effect van het geneesmiddel op de QTc verlenging onafhankelijk van andere factoren te karakteriseren. De parameter die dit effect kwantificeert is echter geen onafhankelijke maat van het risico dat verbonden kan worden aan de verdere ontwikkeling van het molecuul of aan het uiteindelijke gebruik van het geneesmiddel. De waarde van deze parameter is daarentegen een directe afspiegeling van het verband tussen farmacokinetiek en farmacodynamiek die, samen met andere bronnen van variabiliteit, het effect bepaalt.

Daarnaast maakt deze benadering de prospectieve evaluatie van kandidaatgeneesmiddelen mogelijk. Omdat geneesmiddel-specifieke effecten worden gescheiden van de zogenoemde systeem-specifieke effecten kunnen zij ook onder gecompliceerde omstandigheden vastgesteld worden, bijvoorbeeld wanneer het effect op het QT-interval vertraagd optreedt of door een metaboliet wordt veroorzaakt. De voorgestelde Bayesiaanse analyse levert klinisch interpreteerbare effectmaten op die van nut kunnen zijn bij besluitvorming in diverse stadia van de ontwikkeling van een farmacon. De uit deze analysemethode voortvloeiende a-posteriori-verdeling kan direct gebruikt worden voor de vertaling van QTc intervalverlenging naar risico over een reeks van doseringen of concentraties.

De bovengenoemde eigenschappen kunnen relevant zijn in het kader van regelgeving, klinische geneeskunde en geneesmiddelonderzoek. In dit onderzoek werd deze methode gebruikt om voor een bepaalde blootstelling of concentratiebereik de waarschijnlijkheid van een QTc verlenging $\geq$10 milliseconden te bepalen. Deze drempelwaarde kan aangepast worden al naar gelang het te onderzoeken scenario, welk bijvoorbeeld de vergelijking van effecten tussen diersoorten kan betreffen. De essentie van dit deel van het onderzoek is dat een nauwkeurige bepaling van het risico verbonden aan een nieuw kandidaat-molecuul niet mogelijk is zonder geneesmiddel- en systeem-specifieke parameters van elkaar te onderscheiden.

13. 3 TRANSLATIONELE FARMACOLOGIE

13. 3. 1 VAN PREKLINISCHE ONDERZOEKSPROTOCOLLEN NAAR KLINISCHE STUDIES

De huidige *in vitro* experimenten zijn beperkt in hun vermogen om de effecten van therapeutische geneesmiddelconcentraties kwantitatief te voorspellen. De beslissingscriteria die gehanteerd worden bij de ontwikkeling van een kandidaat-molecuul zijn gebaseerd op een kwalitatieve inschatting van het risico op hartritme-stoornissen zonder dat daarin de mate van blootstelling aan de stof meegenomen wordt. Deze tekortkoming kan een relatief hoog percentage vals-positieve uitkomsten tot gevolg hebben. De consequentie hiervan is dat effectieve farmaca verloren gaan voor verdere ontwikkeling. Richtlijn ICH S7B komt in dit bezwaar tegemoet aangezien het stelt dat, om tot een risico-inschatting te komen, naast *in vitro* experimenten ook het geneesmiddeleffect op QT interval *in vivo* moet worden bepaald. Niettemin blijft de bruikbaarheid van deze niet-klinische tests onduidelijk. In principe, dienen de uitkomstmaten voor niet-klinische studies geijkt te worden tegen klinische data. Het gebruik van concentratie-effect relaties als maatstaf hierbij is een nog niet eerder beproefde methode. Het eerder beschreven en op klinische data getoetste Bayesiaanse hierarchische model werd in **hoofdstuk 4** toegepast voor de beschrijving van de relatie tussen geneesmiddelconcentratie en QT verlenging na toediening van cisapride, *d,l*-sotalol en moxifloxacine aan honden. De resultaten van de klinische data analyse en die in honden werden vervolgens aan elkaar gecorreleerd. Doordat het model onderscheid maakt tussen geneesmiddel-specifieke en systeem- (c.q. organisme-)specifieke parameters kon op basis van de analyse in honden een uitspraak gedaan worden over de te verwachten geneesmiddel-geïnduceerde QT verlenging in mensen. De geneesmiddelspecifieke parameter in het model is niet op zichzelf indicatief voor het klinische risico geassocieerd met het gebruik daarvan of het stopzetten van de ontwikkeling van het desbetreffende molecuul. Deze parameter beschrijft uitsluitend de relatie tussen geneesmiddelconcentratie en farmacologisch effect (in dit geval een veiligheidssignaal).

De door ons voorgestelde benadering lijkt een van de belangrijkste beperkingen van eerdere pogingen tot een modelmatige aanpak te overwinnen. Deze nieuwe benadering is in staat geneesmiddeleffecten te

identificeren en is daarom geschikt voor de prospectieve evaluatie van geneesmiddelen in de ontwikkelingsfase. In tegenstelling tot recente voorbeelden van het gebruik van PKPD modellen voor de analyse van QTc interval verlenging [12-15], was ons model in staat preklinisch en klinisch geobserveerde effecten aan elkaar te correleren. Bovendien laat de data analyse voor de genoemde referentiegeneesmiddelen zien dat de beschikbaarheid van individuele farmacokinetische informatie (blootstellingswaarden) in een therapeutische orde van grootte een van de belangrijkste voorwaarden is voor de nauwkeurige vertaling van preklinische effecten naar de mens. Vertalingen op basis van zeer hoge doseringen zoals gebruikelijk bij toxicologie experimenten kunnen leiden tot inaccurate voorspellingen van de grootte van het te verwachten QT effect.

13. 3. 2 VAN GECONTROLEERDE KLINISCHE STUDIES NAAR DE MEDISCHE praktijk

De bevindingen gepresenteerd in dit proefschrift tonen aan dat de vertaalslag niet alleen noodzakelijk is bij de overgang van de preklinische naar de vroege klinische fase van de ontwikkeling van een nieuw molecuul, maar ook wanneer een geneesmiddel toegediend wordt aan de patientenpopulatie tijdens fase II, III en zelfs na registratie (fase IV). De doelstelling in **hoofdstukken 5, 6 en 7** was een analyse van de verschillen tussen de verwachte en de in de medische praktijk waargenomen QTc effecten van geneesmiddelen. Verder belichtten deze hoofdstukken een aantal ethische en praktische vraagstukken met betrekking tot het vergaren van geneesmiddelveiligheidsdata, met name wanneer dit infrequent voorkomende nevenwerkingen betreft. Dergelijke informatie kan enkel gewonnen worden uit epidemiologische studies, welke uitsluitend na registratie van een geneesmiddel plaatsvinden. **Hoofdstuk 5** toont aan dat de criteria die gehanteerd worden bij de selectie van proefpersonen en patienten in klinische studies een verklaring kunnen vormen voor het verschil dat bestaat tussen de bevindingen uit gecontroleerde studies en de relatieve bijdrage van een geneesmiddel aan het waargenomen effect in de doelgroep. Meestal worden met QTc verlenging gepaard gaande storingen in het hartritme van een patient toegeschreven aan het geneesmiddel zonder rekening te houden met andere intrinsieke en extrinsieke factoren die aan het effect kunnen bijdragen. Dankzij het model beschreven in dit proefschrift kon het geneesmiddelspecifieke QTc-verlengingseffect van *d,l*-sotalol in

patienten beschreven worden op een manier die de bijdrage van verschillende bronnen van variabiliteit in acht neemt. De resultaten tonen aan dat de waargenomen QTc waarden in deze populatie niet enkel door het effect van het geneesmiddel verklaard kunnen worden. Een logistische regressie analyse op de patientendata liet zien dat de aanwezigheid van of een geschiedenis van diabetes, hartfalen, hoge bloeddruk en hartinfarct risicofactoren zijn voor QTc verlenging. Verder werd een statistisch significant effect gevonden voor comedicatie en comorbiditeiten die een belangrijke rol spelen op de hartgeleiding en waardoor er een grotere kans bestaat op verhoogde QTc waarden in de data.

In dit hoofdstuk werd gedemonstreerd hoe klinische en epidemiologische data op een geintegreerde manier gebruikt kunnen worden voor de detectie van veiligheidssignalen en voor de verbetering van risicobeheersing van geneesmiddelen. De toepassing van farmacokinetische-farmacodynamische concepten op epidemiologische data liet zien dat de concentratie-effect relatie in proefpersonen (gezonde vrijwilligers), die betrokken worden tijdens de vroege fase van klinische ontwikkeling, van beperkte waarde is aangezien de invloed van factoren die van belang zijn in de doel patientenpopulatie niet meegenomen worden in de analyse. Deze bevindingen extrapolerend naar het domein van TQT studies, kan men vraagtekens zetten bij de validiteit van dit soort studies die de huidige standaard vormen voor de vaststelling van het effect van geneesmiddelen op het hartritme. Indirect wijzen deze bevindingen ook op de tekortkomingen van de zogenaamde double-delta methode voor de analyse van gegevens uit TQT studies. Het gebruik van tijdsgewijs berekende verschillen tussen de behandeling met geneesmiddel en placebo levert resultaten op die niet representatief zijn voor de mogelijk in de praktijk te verwachten geneesmiddel-geïnduceerde QTc verlengingen. Een ander voordeel van het farmacokinetisch-farmacodynamische model is dat zelfs wanneer farmacokinetische parameters bepaald zijn op basis van subtherapeutische geneesmiddelconcentraties, de effecten van blootstelling aan therapeutische concentraties toch voorspeld kunnen worden. Zodoende kunnen uitspraken gedaan worden over de verwachte QTc effecten van klinische relevante doseringen.

Onze benadering heeft ook een onvolkomenheid aan het licht gebracht met betrekking tot de aanname dat bevindingen in een klinische studie extrapoleerbaar zijn naar de doel patientenpopulatie in de geneeskundige

praktijk. De voorwaarden voor de evaluatie van geneesmiddelgerelateerde QTc interval verlenging zijn het onderwerp van **hoofdstuk 6**. Hier werden de gevonden verschillen tussen populaties in klinische studies en die in de praktijk overbrugd en werd het aandeel van een aantal factoren in het QTc effect ingeschat met behulp van simulatiemethodes. De zogenaamde Not-In-Trial Simulatie (NITS) methode opereert onder de aanname dat het QTc interval in de algemene doel populatie een optelsom is van het leeftijdseffect, het geneesmiddeleffect en de effecten van comedicatie en comorbiditeit. Verwacht wordt dat deze methode ook van belang kan zijn voor de inschatting van het QTc effect in risicopopulaties. Dit vereist dat patienten behorend tot deze subpopulaties opgenomen worden in simulatiescenarios. Deze benadering kan nauwkeurigere schattingen leveren met betrekking tot het cardiovasculaire risico en kan oplossingen aanreiken om risico's te vermijden. Zodoende kan een mogelijke terugtrekking van een geneesmiddelregistratie voorkomen worden.

Dit onderzoek belichtte ook methodologische aspecten die verschillen kunnen verklaren in de detectie van veiligheidsignalen zoals die wordt vastgesteld voor en na de toelating van een geneesmiddel tot de markt. Ten eerste is de opzet van veiligheidsstudies sterk gericht op effectiviteit, wat invloed heeft op de gemeten effecten. Ten tweede is het totale aantal patienten dat gedurende klinische studies aan het geneesmiddel wordt blootgesteld niet groot genoeg om een relatief weinig voorkomend maar mogelijk fataal risico op te sporen. Ten derde worden veel subpopulaties met een hoger risico op hartritmestoornis geweerd uit klinische studies. Onder deze subpopulaties kunnen vallen: vrouwen, ouderen, personen die door erfelijkheid of leefpatroon belast zijn met een risico op hart- of andere ziekten geassocieerd met een verminderde repolarisatiereserve (en daardoor meer geneigd zijn QTc interval verlenging te vertonen), personen met een genetische polymorfisme of defect, waardoor de activiteit van geneesmiddel-metaboliserende enzymen of de aangrijpingspunten een ionkanaal aanzienlijk veranderd zijn; personen met een sterkere neiging tot het ontwikkelen van een abnormaal trage hartslag (bradycardie), of elektrolytische onbalans, of gebruikers van geneesmiddelen waarvan bekend is dat ze farmacokinetische of farmacodynamische interacties kunnen vertonen. Verder toonden de in dit proefschrift vermelde resultaten een effect van leeftijd op het onder normale omstandigheden gemeten QTc interval. Deze correlatie verklaart voor een belangrijk deel de waargenomen QTc waarden in de patienten doelgroep, ongeacht van de behandeling.

Om de robuustheid van de bevindingen afgeleid uit de Not-in-Trial simulaties te testen werden data van cisapride als voorbeeld van een tweede geneesmiddel met QTc-verlengingseffect geanalyseerd in **hoofdstuk 7**. In plaats van farmacovigilantieonderzoek te beperken tot meta-analyses van bestaande data (zowel uit gerandomiseerde klinische studies als uit de medische praktijk) laat dit onderzoek zien dat een modelmatige aanpak ook prospectieve "evidence synthesis" mogelijk maakt. Het gebruik van hierarchische modellen stelt ons in staat vaste factoren (effecten van het geneesmiddel of het ziekteproces) en stochastische factoren (variabele effecten, onzekerheden in het model) te integreren. Parameters uit het model worden vervolgens gebruikt om nieuwe scenarios te onderzoeken die klinische relevant kunnen zijn. De QTc metingen in patienten die cisapride toegediend kregen, werden nauwkeurig gekarakteriseerd, in het bijzonder de zeer hoge waarden.

De voorgaande resultaten moeten beschouwd worden als een externe validatie voor het concept van Not-in-Trial Simulatie als een algemene methode voor signaaldetectie en risicobeheersing in de klinische fase van geneesmiddelontwikkeling. De QTc waarden waargenomen in de praktijk voor twee geneesmiddelen met verschillende werkingsmechanismen en therapeutische toepassingen konden correct voorspeld worden. Van twee belangrijke factoren die de bepalend zijn voor het effect werd vastgesteld dat ze onafhankelijk zijn van het toegediende geneesmiddel. Deze factoren zijn leeftijd en comorbiditeit. Hoewel de toepassing van het Not-in-Trial Simulatie concept hier beperkt bleef tot QTc interval verlenging, wordt verwacht dat het ingezet kan worden voor elk neveneffect. Het kan ook gebruikt worden om onderscheid te maken tussen factoren wanneer deze interactie vertonen: met name wanneer het een interactie tussen twee geneesmiddelen of tussen een geneesmiddel en het ziekteproces betreft en de interactie specifiek is voor een bepaald geneesmiddel.

De Not-in-Trial Simulatie methode biedt voordelen ten opzichte van zogenaamde meta-analyses van bestande studies, voor welke het resulterende model vooral een grote, samengestelde data set beschrijft. Met behulp van de simulatiemethode kunnen onderzoekers hypothetische populaties verkennen en verschillende doseringen testen die niet beperkt zijn tot de experimentele data.

De voornaamste beperking van de hier voorgestelde methode is het ontbreken van stochastische variabiliteit op het effect van geneesmiddelblootstelling. De inclusie van een dergelijke variabiliteit (welke verschillen tussen patienten verklaart) zou de bepaling en de voorspelbaarheid van de relatie tussen blootstelling en effect verbeteren. Meer onderzoek op basis van nieuwe onderzoeksprotocollen is nodig om de rol van farmacokinetische variabiliteit in farmacovigilantie studies beter te verklaren.

13. 3. 3 DE VOORSPELLENDE WAARDE VAN QTC INTERVAL VERLENGING VOOR CARDIOVASCULAIR RISICO

Het primaire doel van dit onderzoek naar het QT/QTc-verlengingseffect van nieuwe moleculen was het verminderen van voorvallen van geneesmiddelgeïnduceerd hartritmestoornissen (met uitzondering van geneesmiddelen die een gewenst therapeutisch effect hebben op het hartritme). Een doelstelling in het verlengde hiervan was gericht op het verminderen van gevallen van plotselinge hartdood. Dit doel was gebaseerd op het veronderstelde verband tussen verlenging van het QTc interval, hartritmestoornis en plotselinge hartdood. Echter, gegevens die dit verband aantonen waren niet bechikbaar. Daarom werd **in hoofstuk 8** een omvangrijke evaluatie uitgevoerd waarin de voorspellende waarde van QT interval verlenging onderzocht werd in 182 gevallen van plotselinge hartdood (data afkomstig van het Rotterdam Study Cohort). Hier werd vastgesteld dat de mate van optreden van plotselinge hartdood beter gecorreleerd is aan een hoge absolute waarde van het QTc interval dan aan QTc verschillen. Een verlenging van meer dan 10 milliseconden ten opzichte van de basiswaarde was niet van statistisch belang. Deze bevindingen zijn in overeenstemming met resultaten die laten zien dat het relatieve risico op een levensbedreigende hartritmestoornis hoger is bij een hartslag van meer dan 84 slagen per minuut en dat deze stoornissen zelden voorkomen bij een QTc interval korter dan 500 milliseconden.

De hier getoonde resultaten doen twijfel rijzen over het gebruik van QT interval verlenging als maat voor het cardiovasculaire veiligheidsrisico. In feit werd in een analyse van de incidentie van plotselinge hartdood in deze populatie geen correlatie gevonden met de veranderingen in het QT interval of de tijd tussen het laatst opgenomen electrocardiogram en overlijden. Dit betekent dat het gebruik van absolute QTc waarden voor de bepaling van het

risico op plotselinge hartdood betrouwbaarder zou kunnen zijn. Deze bevindingen weerspreken de hypothese die aan de basis staat van de richtlijnen in ICH E14 over de evaluatie van het cardiovasculaire risico van geneesmiddelen. In deze richtlijnen is het veiligheidsrisico gedefinieerd aan de hand van een drempelwaarde voor verlengingen van 10 milliseconden. Het verband tussen het verschil van 10 milliseconden in QTc en plotselinge hartdood werd echter niet statistisch significant bevonden. Wij voorzien daarom aanpassingen aan twee belangrijke concepten in toekomstige revisies van de ICH E14 richtlijnen: de reden voor de keuze van een drempelwaarde van 10 milliseconden en het gebruik van veranderingen van basislijnwaarden van QTc om risico's te bepalen. Uitgaande van onze bevindingen kan geconcludeerd worden dat veranderingen in QTc veroorzaakt door geneesmiddelen het cardiovasculaire risico niet noodzakelijkerwijs verhogen tenzij de QTc basislijnwaarde hoog is.

Hoewel het onmogelijk is een vaste drempelwaarde voor QTc interval op te stellen zonder hierin verdere factoren te betrekken die bijdragen aan het risico op plotselinge hartdood, moet een effectief beleid voor risicobeheersing basislijn-QTc waarden meenemen als een belangrijke voorspellende factor voor hartritmestoornissen. De aanname dat een toename van het QTc interval met meer dan 10 milliseconden het risico op hartritmestoornissen binnen een populatie in gelijke mate verhoogt kon niet worden gerechtvaardigd.

13. 4 DE ROL VAN SIMULATIES IN DE OPTIMALISATIE VAN PROTOCOLLEN VOOR KLINISCHE STUDIES.

13. 4.1 SIMULATIE VAN RR WAARDEN EN QT INTERVALLEN

Omdat de hartslagfrequentie mede de duur van het QT interval bepaalt, moet bij de bepaling van het geneesmiddeleffect op het QT interval voor hartslag gecorrigeerd worden. Het wiskundige verband tussen QT interval en hartfrequentie dient in het model voor QT bepaling opgenomen te worden. Door gebruik te maken van het gecorrigeerde QT interval (QTc) kan geneesmiddelgeïnduceerde effecten onderscheiden worden van de effecten door verschillen in hartfrequentie.

Voor een effectieve toepassing van simulaties als methode voor de bepalingen van QT verlenging is de beschikbaarheid van individuele hartfrequenties over de tijd noodzakelijk. In tegenstelling tot methoden die ruwe data hanteren om verbanden tussen variabelen af te spiegelen (de zgn. "resampling" methoden), werd in **hoofdstuk 9** het verband tussen QT interval en hartfrequentie (RR) over een periode van 24 uur expliciet beschreven met een model. Deze benadering stelt de onderzoeker in staat realistische hartfrequentiereeksen te genereren die vervolgens gebruikt worden als invoerdata voor het model dat het verband tussen QT en RR beschrijft. Aan de hand van de nieuwe methode werden QT en RR metingen gesimuleerd voor discrete tijdspunten, zoals vaak uitvoerd in een klinisch protocol. De variantiestructuur en de correlaties in de gesimuleerde data zijn een afspiegeling van de ruis op de data tengevolge van meet- en staalnameprocedures. Deze nieuwe benadering was essentieel voor de evaluatie en optimalisatie van de klinische protocollen in **hoofdstukken 10 en 11**. In deze evaluatie werden variaties in studieopzet en analysemethode onderzocht op de eruit voortvloeiende statistische type I en II fouten.

13. 4. 2 CONCENTRATIE-EFFECT RELATIES VERSUS DE "*DOUBLE-DELTA*" METHODE

In **hoofdstuk 10** werden grondige QT (TQT) studies gesimuleerd onder de aanname van verschillende populatiegroottes en geneesmiddelgeïnduceerde QT effecten. Ook de studieopzet werd gevarieerd: een gekruiste studie met referentiewaarden voor alle kloktijden (staalnames), een gekruiste studie met drie referentiewaarden en parallelle studies werden bestudeerd. De resultaten toonden aan dat een Bayesiaans hierarchisch model geneesmiddelgeïnduceerde QT effecten met een gelijke of hogere specificiteit en sensitiviteit onderscheidt dan de double-delta methode. Deze bevindingen pleiten voor het gebruik van concentratie-effectrelaties in plaats van de double-delta methode voor de analyse van QTc interval verlenging. Zodoende kunnen aanwijzingen voor of tegen een QT verlengend effect vergaard worden met een significant kleiner aantal studiedeelnemers. Aangezien de hoge kosten van TQT studies voornamelijk toegeschreven worden aan de vereiste studiegrootte, biedt dit resultaat een reden om ethische, statistische en financiele aspecten van de bepaling van QT verlengingen opnieuw te beschouwen. Een ander voordeel van het gebruik van een modelmatige benadering is de mogelijkheid om op een andere

manier met het multipliciteitsprobleem om te gaan (multipliciteit onstaat wanneer een effect meerdere keren op verschillende momenten in de loop van de studie vastgesteld moet worden). Het geneesmiddelgeïnduceerde effect wordt beschreven als een continue functie van de geneesmiddelconcentratie in plaats van het op een aantal discrete tijden te schatten. Deze bemoedigende resultaten vormen de bouwstenen voor zowel een alternatief voor de huidige analysemethode voor TQT studies, als een strategie om de kosten van dergelijke studies te reduceren.

Gezien de aanzienlijke winst in sensitiviteit die te boeken valt met een modelmatige benadering voor de bepaling van QT interval verlenging, werd in **hoofdstuk 11** de mogelijkheid onderzocht om het QT verlengingsrisico gedurende de standaard uitgevoerde "first-in-human" (FTIH) studie te bepalen. Men kan ervan uit gaan dat het risico op een QT verlengend effect correct ingeschat kan worden wanneer de resultaten van een dergelijke bepaling beschouwd worden in samenhang met aanwijzingen voor een concentratie-effectrelatie en de daaraan verwante kans op QTc interval verlenging in preklinische diermodellen. Een ander voordeel van de aanpak op basis van modellen is de mogelijkheid om het QT verlengende effect van een geneesmiddel over het gehele therapeutische concentratiebereik te karakteriseren. FTIH studies vormen een verplicht en vastomschreven onderdeel van het klinische ontwikkelingpad. Slechts geringe aanpassingen aan het schema van metingen zijn noodzakelijk om ECG bepalingen geschikt te maken voor een modelmatige analyse van het geneesmidel effect op het QTc interval.

Daarnaast lieten de analyseresultaten zien dat nauwkeurige schattingen van de grootte van het QT verlengende effect gemaakt konden worden zonder gebruik te maken van een positieve controlearm. Het nut of relevantie van bestaande gegevens van een referentiegeneesmidel (ofwel positieve controle) in de vorm van een a priori schatting wordt hiermee echter geenszins ontkend.

13.5 AANBEVELINGEN

Dit proefschrift belicht de gebreken van verscheidene aspecten van de ICH E14 richtlijn. Verder draagt het alternatieven aan en doet het aanbevelingen ter verbetering. In de eerste plaats werd een Bayesiaans model

geïntroduceerd dat het verband beschrijft tussen geneesmiddelconcentratie en QTc interval verlenging. Met dit model kunnen geneesmiddelafhankelijke effecten over een relevant concentratiebereik voorspeld worden. Deze methode verdient de voorkeur boven de puur empirische double-delta methode. Bovendien blijkt het gebruik van doseringen hoger dan de therapeutische niet noodzakelijk te zijn, maar ook indien deze wel toegepast worden, kan het effect van een dosis in het therapeutische bereik eenvoudig herleid worden met het hierarchische model. Verder werd aangetoond dat de Bayesiaanse benadering QTc verlengingen over een tijdscontinuum nauwkeuriger en preciezer voorspelt. Een ander voordeel van een op modelmatige aanpak is de mogelijkheid om QTc effecten te beschrijven en te vergelijken voor een spectrum van farmaca en organismen. De aan deze methode inherente beschikbaarheid van a posteriori waarschijnlijkheidsverdelingen stelt de onderzoeker in staat de kans op een vooraf gedefinieerde verlenging van het QTc interval te bepalen. Door hetzelfde model opnieuw toe te passen voor de analyse van data verkregen uit een FTIH studie met enkele aanpassingen met betrekking tot de ECG opnames, kan de kans dat een farmacon het hartritme beinvloedt vastgesteld worden zonder dat hiervoor een aparte TQT studie uitgevoerd hoeft te worden. Dit houdt in dat besluiten om de ontwikkeling van een geneesmiddel al dan niet voort te zetten eerder genomen kunnen worden.

13.6 VOORUITBLIK

De resultaten gepresenteerd in dit proefschrift onderstrepen de behoefte aan een nieuwe denkwijze bij de vaststelling van de cardiovasculaire veiligheid van geneesmiddelen. In deze denkwijze is de “evidence synthesis” niet langer ondergeschikt aan het genereren van nieuwe gegevens. De resultaten bevestigen verder de noodzaak om de huidige richtlijnen te herzien. Met name de inhoud van het ICH E14 document blijkt niet in overeenstemming met recente wetenschappelijke inzichten.

Het is van belang dat klinisch onderzoekers, medici in de gezondheidszorg en regel- en wetgevende autoriteiten inzien dat de opzet van een geneesmiddelstudie van invloed is op de gegevens die de studie voortbrengt. Verschillen tussen de onderzoekspopulatie bestudeerd tijdens de ontwikkelingsfase en de patientenpopulatie in de medische praktijk mogen niet genegeerd worden. Intrinsieke en extrinsieke factoren of covariaten

dragen bij aan de tot standkoming van het behandelingsresultaat. De farmacologische effecten van het geneesmiddel worden door deze factoren geantagoneerd, tenietgedaan of anderszins gemodificeerd. Binnen het regelgevende kader zou de vraag moeten rijzen hoe efficient en betrouwbaar klinische studies zijn in het vaststellen van het risico op levensbedreigende hartritmestoornissen, gezien de verschillen tussen populaties, de intrinsieke ruis op de data en het relatief infrequent voorkomen van significante geneesmiddelgerelateerde effecten [16].

Zoals de titel van dit proefschrift suggereert bestaat er een kloof in de vertaling van veiligheidssignalen in klinische studies naar de medische praktijk, welke overbrugd moet worden als onderdeel van de risicobeheersing gedurende de ontwikkeling van een nieuw molecuul. Daarom is het belangrijk tijdens de klinische ontwikkeling de aandacht te vestigen op het concept van “evidence synthesis” en ervoor te zorgen dat het verband tussen farmacokinetiek en farmacodynamiek bekend is. De beschikbaarheid van op modellen gebaseerde analyse- en simulatiemethodes vormen de essentiele gereedschappen om deze kloof te dichten. Wij verwachten dat vooruitgang geboekt kan worden in de voorspelling van cardiovasculaire geneesmiddelveiligheid door de toepassing van het concept van Not-in-Trial simulaties uit te breiden naar meerdere eindpunten en door mechanistische modellen te combineren met de voorgestelde Bayesiaanse hierarchische benadering. Gezien de toenemende beschikbaarheid van data over geneesmiddelveiligheid afkomstig uit observationele studies zal de toegang tot informatie niet langer de beperkende factor zijn in de overbrugging van eerdergenoemde kloof. Onafhankelijk van de data blijft het verder van kritiek belang dat in de modelmatige benadering onderscheid gemaakt kan worden tussen geneesmiddelspecifieke en systeemspecifieke parameters.

Het concept van de Not-in-Trial simulaties, ontwikkeld in dit proefschrift, kan een verdere basis bieden voor de integratie van klinische en farmaco-epidemiologische data. Het werd in eerste instantie geïntroduceerd als een instrument om de uitwerking te bepalen van de strikte inclusie- en exclusiecriteria die opgesteld worden voor klinische studie-populaties, alsook om de invloed van de relatief korte duur van klinische studies na te gaan. Deze beperkingen dragen ertoe bij dat infrequent of zeldzaam voorkomende veiligheidssignalen pas lange tijd nadat het geneesmiddel op de markt beschikbaar komt gedetecteerd worden. De mogelijkheid om

dergelijke signalen al vóór de registratie te anticiperen biedt de mogelijkheid om risicoverminderende maatregelen te onderzoeken en implementeren.

Het idee achter Not-in-Trial simulaties is om relevante fysiologische parameters en geneesmiddel-specifieke parameters te combineren in een model en vervolgens de invloed van andere factoren zoals comborbiditeit en comedicatie op de parameterwaarden te bepalen. In één van zijn publicaties benadrukte Black " dat het onterechte conflict tussen diegenen die onder alle omstandigheden voor gerandomiseerde studies pleiten en diegenen die geloven dat observationele data afdoende zijn, plaats moet worden gemaakt voor wederzijds begrip voor de complementaire rol van de twee benaderingen." [17].

Regelgevers en klinisch onderzoekers vertrouwen bij de evaluatie van een geneesmiddel onder meer op een positieve waarde van de ratio tussen therapeutische waarde en het desbetreffende risico. Hierbij zijn zij afhankelijk van data die gegenereerd wordt in klinische of observationele studies [18-21]. In de praktijk zijn dit vaak langdurige klinische fase IV studies of, recentelijk, observaties gemaakt in het kader van een voorwaardelijke registratieverstrekking ("progressive licensing"). Deze oplossing legt de last uiteindelijk bij patienten en geneesmiddelontwikkelaars, met het risico dat bewijs van onveiligheid pas verkregen wordt nadat een nevenwerking is opgetreden in de medische praktijk. Tot dusver zijn pogingen om alternatieven te ontwikkelen voor formele prospectieve studies beperkt gebleven tot "retrospectieve" meta-analyses van bestaande data uit de klinische praktijk alsook uit gerandomiseeerde klinische trials. Het is essentieel dat meer onderzoek gedaan wordt naar de ontwikkeling van voorspellende methoden en vooruitblikkende strategieen.

De kracht van de Bayesiaanse benadering is beperkt tot de analytische capaciteit om QTc interval verlenging op te sporen en helpt daarom enkel bij de bepaling van de arythmogeniciteit. Zij biedt echter geen oplossing voor de praktische of statistische problemen met betrekking tot de verwerking van ECG metingen. Mogelijke onderwerpen waarin vooruitgang geboekt kan worden zijn het algoritme voor het aflezen van ECG bepalingen, de techniek voor de correctie van meetonnauwkeurigheden, de invloed van de aard van comorbiditeit op de ECG waarden en verscheidene andere covariaten die hun weerslag hebben op deze metingen.

Er wordt veel onderzoek verricht om de precisie van QT metingen te verbeteren. Onder de onderzochte methoden valt het gebruik van geautomatiseerde afleestechnieken. Verder wordt getracht statistische problemen te verhelpen door oorzaken van variabiliteit te achterhalen, nieuwe algoritmes te ontwikkelen en de reproduceerbaarheid van bepalingen te vergroten. In 2001 ontwikkelde de FDA het digitale ECG Warehouse voor de opslag van ECG metingen die aangeleverd worden ten behoeve van registratieverzoeken. Naast zijn primaire functie van opslagfaciliteit vormt het Warehouse een geschikte bron voor onderzoek door universiteiten en bedrijven. Het is een waardevol instrument dat gebruikt kan worden om nieuwe biomarkers voor hartritmestoornissen te bestuderen, deze te correleren aan klinische en biochemische maatstaven voor cardiovasculaire veiligheid, om hiermee de diagnose en behandeling van cardiovasculaire aandoeningen te verbeteren. Verder biedt het de mogelijkheid variabiliteit in fysiologische variabelen te bepalen en deze te gebruiken om a priori verdelingen op te stellen voor toekomstige Bayesiaanse analyses van QTc interval verlengingen op basis van farmacokinetische-farmacodynamische relaties.

Geneesmiddelonderzoekers moeten “oppassen voor de kloof” tussen gecontroleerde klinische trials en de geneeskundige praktijk. De tijd is rijp om deze problematiek aan te pakken door de verschillen te kwantificeren, te voorspellen, te interpreteren en uiteindelijk te overbruggen. Het kan niet genoeg benadrukt worden dat het nu mogelijk is een op modellen en simulaties gebaseerde benadering voor risicobeheersing te ontwikkelen.

REFERENTIES

[1] Kowey PR, Malik, M. The QT interval as it relates to the safety of non-cardiac drugs. European Heart Journal Supplements 2007;9(Supplement G).

[2] Letsas KP, Letsas, G. P. Drug-Induced Proarrhythmia: QT Interval Prolongation and Torsade de Pointes. Hospital Chronicles 2011;6(3):118-22.

[3] Piccini JP, Whellan DJ, Berridge BR, Finkle JK, Pettit SD, Stockbridge N, et al. Current challenges in the evaluation of cardiac safety during drug development: translational medicine meets the Critical Path Initiative. Am Heart J 2009 Sep;158(3):317-26.

[4] Turner JR. Drug safety, medication safety, patient safety: An overview of recent FDA guidances and initiatives. Regulatory Rapporteur 2009;6(4).

[5] Pollard CE, Valentin JP, Hammond TG. Strategies to reduce the risk of drug-induced QT interval prolongation: a pharmaceutical company perspective. Br J Pharmacol 2008 Aug;154(7):1538-43.

[6] Anon. ICH E14 The clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs. 2005.

[7] Anon. ICH Topic S7A Safety Pharmacolog Studies for Human Pharmaceuticals. 2000.

[8] Anon. ICH Topic S7B The nonclinical Evaluation of the Potential for delayed Ventricular Repolarization (QT Interval Prolongation) by Human Pharmaceuticals. 2005.

[9] Pater C. Methodological considerations in the design of trials for safety assessment of new drugs and chemical entities. Curr Control Trials Cardiovasc Med 2005 Feb 3;6(1):1.

[10] Turner JR. Analyzing Safety Data. Key Statistical Concepts in Clinical Trials for Pharma: Springer; 2012. p. 11-21.

[11] Pepe M, Longton G, Janes H. Estimation and Comparison of Receiver Operating Characteristic Curves. Stata J 2009 Mar 1;9(1):1.

[12] Cavero I. Using pharmacokinetic/pharmacodynamic modelling in safety pharmacology to better define safety margins: a regional workshop of the Safety Pharmacology Society. Expert Opin Drug Saf 2007 Jul;6(4):465-71.

[13] Fleury A, Lave T, Jonsson F, Schmitt M, Hirkaler G, Polonchuk L, et al. A pharmacokinetic-pharmacodynamic model for cardiovascular safety assessment of R1551. J Pharmacol Toxicol Methods 2010 Jan-Feb;63(1):123-33.

[14] Langdon G, Davis JD, McFadyen LM, Dewhurst M, Brunton NS, Rawal JK, et al. Translational pharmacokinetic-pharmacodynamic modelling; application to cardiovascular safety data for PF-00821385, a novel HIV agent. Br J Clin Pharmacol 2010 Apr;69(4):336-45.

[15] Ollerstam A, Visser SA, Duker G, Forsberg T, Persson AH, Nilsson LB, et al. Comparison of the QT interval response during sinus and paced rhythm in conscious and anesthetized beagle dogs. J Pharmacol Toxicol Methods 2007 Sep-Oct;56(2):131-44.

[16] Bonate PL, Russell T. Assessment of QTc prolongation for non-cardiac-related drugs from a drug development perspective. J Clin Pharmacol 1999 Apr;39(4):349-58.

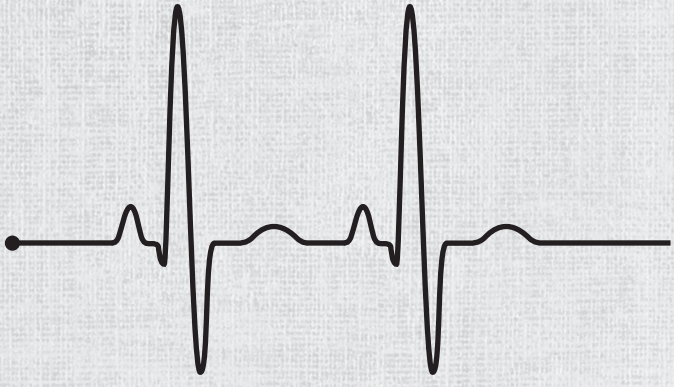
[17] Black N. Why we need observational studies to evaluate the effectiveness of health care. BMJ 1996 May 11;312(7040):1215-8.

[18] Baker SG, Cook NR, Vickers A, Kramer BS. Using relative utility curves to evaluate risk prediction. J R Stat Soc Ser A Stat Soc 2009 Oct 1;172(4):729-48.

[19] Vickers AJ, Cronin AM, Kattan MW, Gonen M, Scardino PT, Milowsky MI, et al. Clinical benefits of a multivariate prediction model for bladder cancer: a decision analytic approach. Cancer 2009 Dec 1;115(23):5460-9.

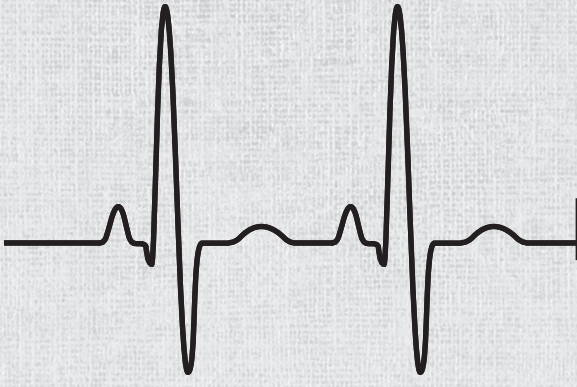
[20] CIOMS. Benefit-Risk Balance for Marketed Drugs: Evaluating Safety Signals. Geneva: CIOMS Working Group V; 1998.

[21] Walker S, Philips, L., Cone, M. Benefit-Risk Assessment Model for Medicines: Developing a Structured Approach to Decision Making. Washington DC: CMR International Institute for Regulatory Science; 2005.



"Now this is not the end. It is not even the beginning
of the end. But it is, perhaps, the end of the beginning."

Winston Churchill



EPILOGUE

ACKNOWLEDGEMENTS

Forrest Gump's momma always said "*Life* was *like a box of chocolates. You never know what you're gonna get.*" My life has certainly always been like a box of chocolate, especially during the past few years! Well, first I didn't always know I was going to go for a PhD and second, the journey has definitely been filled with both sweet and bittersweet moments.

I most certainly have to begin by acknowledging those who believed in me before I even realized this is what I wanted. Oscar, in the almost-ten years that I have known you, you've never failed to encourage me, to push me to my limits and beyond. You have not only expanded my world scientifically, but also culturally and gastronomically. You knew I would return to Leiden before I could imagine doing a second degree. I am truly lucky to have a champion like you on my side. But of course, this would not have been possible without the introduction from Meindert. Meindert, I will never forget the day I saw your picture on the Leiden University website. Albeit an unusual way to mark the start of one's career, it's unlike sampling new chocolate based on attractive packaging. You have always given me the pivotal guidance and advice at the time of need. Your calm and no-nonsense mannerism have kept me focused and on track all the way to the finish line. And how can I go on without mentioning Dymphy, who was brave enough to take on the challenge to work with a rookie who knew absolutely nothing about modelling but a few ordinary differential equations.

Everyone at the division of Pharmacology has also made the experience that much sweeter. All the staff, post-docs and PhD colleagues have all been welcoming and accommodating. Much fond memories are made in and outside of the office with Rob, Liesbeth, Bert, Catherijne, Stina, Stephan, Massoud, Donato, Charlotte, Flora, Margot, Geert, Jaap, Maurice, Amit, Vincenzo, Vincent, Joost, Jasper, Chiara, Tarj, Roos, Massimo, Ashley, Francesco, Sven, Chantal, Imke, Marta, Erik, Bea, Linda, Iba and Elke. I will never forget our endless discussions on pharmacology, R, NONMEM, Dutch weather, Gorlaeus cafeteria, karnemelk, kroket and football. I am also grateful for having wonderful students, Francesco and Morris, whom have worked so hard to help me execute a tremendously huge volume of simulations.

To be able to belong to another equally fun and stimulating group at the department of informatics in Erasmus Medical Centre has been nothing less than awesome. Miriam, I can still remember your glamorous entrance the first time I met you at the top of the stairs when you visited Leiden. You have been so supportive during my first encounters in the field of epidemiology and have continued to open doors to many opportunities and stuck by me throughout the last four years. Jeanne, I am very appreciative for your insights and patience. I have enjoyed our many discussions immensely. I am also grateful to belong to the same group as Jan, Ewout, Kris, Mees, Marcel, Carmen O., Tineke, Desiree, Rene, Bruno and proud to call Inge, Fatma, Roelof, Seppe, Eva, Emine, Alessandro, Ana, Ann, Vera, Sandra, Jelmer, Carmen F., Gianluca, Martijn, Silvana, Marius, Katia, Mendel, Charlotte, Eva van S., Eline, Toke, Rikje, Esben, Nico, Leo and Preci my friends.

Of course my project would not have been possible without the generosity from the Top Institute Pharma. Apart from the financial support, it has provided invaluable opportunities to collaborate with and learn from veterans in the field of pharmacology. I recognise the valuable interactions with Eliane Fuseau, Mats Karlsson, Andrew Hooker, David Spiegelhalter, David Lunn, Bronagh Heath, Nick McMahon, Jackie Bloomer, Piet Van Der Graaf, Mark Holbrook, John Davis, Dinesh de Alwis, Ivelina Gueorguieva, Charles Benson, Derek Leishman, An Vermeulen, David Gallacher, Rik de Greef, Thomas Kerbusch, Sandra Visser, Gezim Lahu, Axel Facius and Kevin Krudys who so kindly shared his expertise and his post-doctoral work as a starting point for this thesis.

My time in the Netherlands has also been enriched by Maren, Richie, Chrissy, Oli, Anna, Sandro, Nicole, Mike, Thomas, Andrew and David. Thank you for welcoming me into your lives and providing distractions for my "hard" time as a PhD student in Holland.

Having spent a combined two years as an undergraduate trainee and subsequently a graduate fellow at GlaxoSmithKline in London, I had the tremendous privilege to gain knowledge from and have fun with Abi, Pam, Isabelle, Daren, Bart, Gailing, Misba, Ann, Kevin S., Alex, Shuying, Sabrina, Claire, Guy, Alienor, Jonathan, Sophie, Alun, Stefano, Bela, Nicolas, Nadia, Chris, Georgios and Lorenzo. Hugo, who knew we would meet again under such different circumstances and in a different country. I will always be inspired by your constant attention to details and impeccable self-discipline.

Thank you for coming to my rescue and saved me from writing the Dutch Summary. Uncle Chao, your wisdom and constant life lessons have made me a better person today. Lutz, I shall never take for granted the foundation you have bestowed upon me and I can only hope to be able to continuously learn from you for many more years to come.

London has been much more than a learning experience. It has become a home away from home for the last few years with the generosity from Auntie Tessie, Norman, Daniel, Laura and Steve - you are my second family. My time there has also been filled with lasting memories, courtesy of Auntie Mary, Bryn, Auntie Sonia, Steve, Eva, Rashit, Nicole, Alan, Kevin, Claudia and Amy.

Being away from Toronto in the past few years has had its drawbacks. But I am very fortunate to have friends like Alyssa, Tom, Anthea, Mike, Anson, Jen, Ernest, Angela, Jeff, Luka, Karen, and Omar who never failed to satisfy my continuous cravings for sushi. Omar, a special note to you for all the design work you have put into this book and entertaining my relentless pursuit of nitty-gritty details - the colour grey will never be the same again.

My paranymphs, Elke and Preci - I cannot imagine surviving the last few years without you or share hotel rooms with anyone else around the world. Elke, you helped me navigate my way through life in Leiden and you were always the voice of reason when I was frustrated. Preci, my crazy twin sister, you always had a way to make me laugh and drive me crazy like nobody else can.

Auntie Meg, you have always reminded me that I can do whatever I want if I set my mind to it. I have become the self-reliant person I am today largely because of you. Mommy and daddy, you have given me a life full of opportunities and experiences that are unparalleled. You have trained me to always strive for excellence and never settle for mediocrity. I will always be grateful for the sacrifices you have made along the way and will forever be proud to be your daughter.

Finally, a big huge thank you to all those who I may have forgotten to mention. This journey would not have been as amazing without your friendships and unconditional support :)

PHD PORTFOLIO

Name	Anne S. Y. Chain
University	Leiden University, Division of Pharmacology
PhD Period	November 2007 - December 2011
Promotors	Prof. dr. Meindert Danhof, Prof. dr. Miriam C.J.M. Sturkenboom
Co-promotor	Dr. Oscar Della Pasqua

1. PHD TRAINING

Research Skills

2008	Advanced epidemiology course, Erasmus Winter Program
2008	Bayesian analysis, Reading University, UK
2008	ECG workshop, M&K courses, UK
2008	Six concepts to statistical thinking, Royal Statistical Society, UK
2008	Bayes Course, Cambridge University, UK
2009	Pharmacokinetic-pharmacodynamic modelling of continuous and categorical data in NONMEM, Uppsala Group, Washington DC, USA
2009	Binary and categorical data analysis, Leiden, The Netherlands
2010	Advanced R Course, Mango Solutions, Leiden, The Netherlands
2010	Drug Delivery Cycle, TI Pharma, The Netherlands

2010 Business and Entrepreneurial Skills,
TI Pharma, The Netherlands

2011 R Graphics Course, Mango Solutions, San Diego, CA, USA

Oral Presentations

2009 *Not-In-Trial Simulation: Predicting Cardiovascular Risk from Clinical Trial Data.* European Association of Clinical Pharmacology and Therapeutics Meeting 2009, Edinburgh, United Kingdom.

2010 *Translation of Cardiovascular Risk in Drug Development.* 6th International Symposium on Measurements and Kinetics of In Vivo Drug Effects, Noordwijkerhout, The Netherlands.

2010 *Translating Cardiovascular Risk from Clinical Trial Data.* International Conference on Pharmacoepidemiology and Therapeutic Risk Management 2010, Brighton, United Kingdom.

Poster Presentations

2008 *Bridging Cardiovascular Risk from Clinical Trials to Real Life Population.* Population Approach Group Europe Meeting 2008, Marseille, France.

2009 *Not-In-Trial Simulation: Predicting Cardiovascular Risk from Clinical Trial Data.* Population Approach Group Europe Meeting 2009, St. Petersburg, Russia.

2010 *Translation of ECG Findings into Sudden Cardiac Death.* 6th International Symposium on Measurements and Kinetics of In Vivo Drug Effects, Noordwijkerhout, The Netherlands.

2010 *QTc Prolongation Liability: Prospective Use of Not-In-Trial Simulations.* International Conference on Pharmaco-epidemiology and Therapeutic Risk Management 2010, Brighton, United Kingdom.

2011 *Generating RR Intervals for the Assessment of Drug Effects in Clinical Trial Simulations.*
American Conference on Pharmacometrics 2011, San Diego, CA, USA.

2011 *Can First-Time-In-Human Trials Replace Thorough QT Studies?*
Population Approach Group Europe Meeting 2011, Athens, Greece.

2011 *Relevance of QT-RR correlations in the assessment of QTc-interval prolongation in clinical trial simulations.*
Population Approach Group Europe Meeting 2011, Athens, Greece.

2011 *Validation of a Not-In-Trial Simulation Tool for Cardiovascular Risk Mitigation.*
International Conference on Pharmacoepidemiology and Therapeutic Risk Management 2011, Chicago, IL, USA.

International Conferences

2008 17th Population Approach Group Europe Meeting (PAGE) 18-20 June, 2008, Marseille, France.

2009 18th Population Approach Group Europe Meeting (PAGE) 23-26 June, 2009, St. Petersburg, Russia.

2009 9th European Association of Clinical Pharmacology and Therapeutics Meeting (EACPT). 12-15 July, 2009, Edinburgh, United Kingdom.

2009 25th International Conference on Pharmacoepidemiology and Therapeutic Risk Management (ICPE). 16-19 August, Providence, Rhode Island, USA.

2010 6th International Symposium on Measurements and Kinetics of In Vivo Drug Effects. 21-24 April, 2010, Noordwijkerhout, The Netherlands

2010 19th Population Approach Group Europe Meeting (PAGE) 8-11 June, 2010, Berlin, Germany.

2010 26th International Conference on Pharmacoepidemiology and Therapeutic Risk Management (ICPE). 9-22 August, 2010, Brighton, United Kingdom.

2011 American Conference on Pharmacometrics (ACoP). 3-6 April, 2011, San Diego, California, USA.

2011 20th Population Approach Group Europe Meeting (PAGE) 7-10 June, 2011, Athens, Greece.

2011 27th International Conference on Pharmacoepidemiology and Therapeutic Risk Management (ICPE). 14-17 August, 2011, Chicago, Illinois, USA.

Other Presentations and Seminars

2008 - 2011 Leiden / Amsterdam Centre for Drug Research (LACDR) Spring Symposiums.

2008 - 2011 TI Pharma Bi-annual Meetings.

2008 - 2011 Division of Pharmacology Lunchtime Seminars.

2008 - 2010 Research Colloquia in Medical Informatics, Department of Medical Informatics, Erasmus University Medical Centre.

2009 - 2010 Medical Informatics Day, Erasmus University Medical Centre.

2. TEACHING ACTIVITIES

2007 - 2010 Teaching assistant to undergraduate students in biopharmaceutical sciences.

- Coordinated course logistics and assisted in mock clinical trials
- Lectured on clinical trial designs and protocol writing
- Conducted literature reviews with students

LIST OF PUBLICATIONS

- A.S.Y. Chain, K.M. Krudys, O. Della Pasqua.
  Assessing the Probability of Drug-induced QTC-interval Prolongation During Clinical Drug Development.
  Clinical Pharmacology and Therapeutics, 2011 Dec; 90(6): 867-75.

- A.S.Y. Chain, M.C.J.M. Sturkenboom, M. Danhof, O.E. Della Pasqua.
  Establishing in vitro to clinical correlations in the evaluation of cardiovascular safety pharmacology.
  Drug Discovery Today - Technologies, 2012. [In press]

- A.S.Y. Chain, J.P. Dieleman, C. van Noord, M. Danhof, M.C.J.M. Sturkenboom, O. Della Pasqua.
  Not-In-Trial Simulation I: Bridging Cardiovascular Risk from Clinical Trials to Real Life Conditions.
  British Journal of Clinical Pharmacology, 2012. [In press]

- A.S.Y. Chain, F. Bellanti, M. Danhof, M.C.J.M. Sturkenboom, O. Della Pasqua.
  Simulating QT and RR Intervals for the Evaluation of Drug-Induced QTc-Interval Prolongation.
  British Journal of Clinical Pharmacology, 2012. [In press]

- D. Huntjens, A. Strougo, A. Chain, A. Metcalf, S. Summerfield, D. Spalding, M. Danhof, O. Della Pasqua.
  Population Pharmacokinetic Modelling of the Enterohepatic Recirculation of Diclofenac and Rofecoxib in Rats
  British Journal of Pharmacology, 2008, 53(5): 1072-8.

ABOUT THE AUTHOR

Anne S. Y. Chain was born on April 26th, 1982 in Hong Kong. Her family immigrated to Toronto, Ontario, Canada when she was 10 years old. In 2001, she completed high school as an Ontario Scholar award winner at North Toronto Collegiate Institute. She began Engineering Science at the University of Toronto in the same year, specializing in Biomedical Engineering in her third year of studies. In the summer of 2003, she was first introduced to pharmacokinetics modelling in a five-week traineeship at the Division of Pharmacology in Leiden University, under the guidance of Professor dr. Meindert Danhof and Dr. Oscar Della Pasqua. With the supervision of Ms. Dymphy Huntjens, she worked on a project involving enterohepatic recirculation of rofecoxib in rats and learnt the basic concepts of population compartmental modelling. Between 2004 and 2005, she continued on an industrial internship in Modelling and Simulation with Dr. Della Pasqua and Dr. Lutz Harnisch at GlaxoSmithKline, in London, United Kingdom as part of her undergraduate degree. While there, she was given the opportunity to evaluate the effect of *d, l*-sotalol on QT/QTc-interval prolongation in healthy volunteers. Anne graduated university with a Bachelors of Applied Science (BASc) degree in 2006.

Soon after graduation, Anne began working in Sapient Corporation, an international business consultancy in Toronto. In November 2007, she left her position as a senior associate in project management and returned to Leiden University to begin her PhD programme in cardiovascular safety. The research project described in this thesis is part of a work package belonging to the PKPD Modelling Platform within the Top Institute Pharma, a tripartite consortium, involving industry, academia and the Dutch government.

In the Spring of 2012, Anne joined the Modelling and Simulation group headed by Dr. Sandra Allerheiligen at Merck & Co. Inc. in Rahway, New Jersey, USA.

LIST OF ABBREVIATIONS

AAI.	Arm ankle index
ADP.	Action potential duration
ADR.	Adverse drug reaction
BMI.	Body mass index
CHMP.	Committee for Human Medicinal Products
CNS.	Central nervous system
CPMP.	Committee for Proprietary Medicinal Products
CI.	Confidence interval
CTS.	Clinical trial simulation
CV.	Coefficient of variance
CVS.	Cardiovascular safety
DIC.	Deviance information criterion
EAD.	Early after-depolarization
ECG.	Electrocardiogram
EMA.	European Medicine Agency
FDA.	Food and Drug Administration
GP.	General practitioners
ICD.	International classification of diseases
ICH.	International conference on harmonization
LBBB.	Left bundle-branch block
LQTS.	Long-QT syndrome
LVH.	Left ventricular hypertrophy
MBDD.	Model-based drug discovery
MCMC.	Markov chain Monte Carlo
MEANS.	Modular ECG analysis system
MMRM.	Mixed model for repeated measure
NCE.	New chemical entities

NPDE.	Normalised prediction distribution error
PKPD.	Pharmacokinetic-pharmacodynamic
QT.	Time interval between the Q and T peak in the electrocardiogram
QTc.	Heart-rate corrected QT-interval
RBBB.	Right bundle-branch block
RR.	1. Relative risk 2. Time interval between successive R peaks in the electrocardiogram
SCD.	Sudden cardiac death
TdP.	Torsade de pointes
TQT.	Thorough-QT
VPC.	Visual predictive check
WHO.	World health organization

STELLINGEN / PROPOSITIONS TO THIS DISSERTATION

1. A major gap in the evaluation of cardiovascular safety is the reliance on empirical evidence. *This thesis*

2. The predictive power of the techniques and methodologies currently available for the evaluation of cardiovascular risk depends upon careful consideration of the clinically relevant conditions. *This thesis*

3. The use of model parameterisation discriminating drug- and system-specific properties eliminates the bias arising from the current regulatory guidelines. *This thesis*

4. The integration of pharmacological concepts with pharmaco-epidemiology offers a unique opportunity to bridge the gap between clinical trials and real life conditions. *This thesis*

5. The concept of Not-In-Trial Simulations defies the hurdles imposed by the evidentiary axiom "the absence of evidence is not evidence of absence". *This thesis*

6. The possibility to communicate QTc prolongation in terms of probability provides transparency that greatly enhances the decision-making processes throughout drug development. *This thesis*

7. The marked differences observed between clinical trials and post-marketing use confirm that pre- and post-approval are two distinct worlds. *(Martin, K. et al., Br J Clin Pharmacol 2003; 57: 86–92)*

8. The most important- and vexing- questions in clinical cardiology today can be phrased quite simply: ""Why did he die on Tuesday and not on Monday?" If we can identify the transient initiating event, we can begin to develop more effective strategies for prevention of SCD. *(Rubart M. & Zipes D.P., JCI 2005;115(9):1–11)*

9. In order to make a fully informed treatment decision, it is logical that all available information be used. *(Lynd L.D. & O'Brien B.J., J Clin Epidemio 2004; 57: 795–803)*

10. The forecast is that complex, mixed-effects, mechanistic PD models combined with Bayesian estimation and incorporating different sources of prior knowledge will provide a paradigm for future model development. *(Csajka C. & Verotta D., JPP 2006; 33: 227-279)*

11. Even a Nobel laureate can be wrong. Today Carrel would not have stated that, "a few observations and much reasoning lead to error; many observations and a little reasoning to truth." *~Alexis Carrel, Nobel Prize winner, 1912.*

12. The phrase "mind the gap" has never been so applicable to the pharmaceutical industry as in the last five years. The gaps in the drug development process are larger than those on any train platform.