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
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 (Ca²⁺) 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, terikalanl, 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).
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.

BAZETT’s FORMULA: $QTcB = \frac{QT}{\sqrt{RR}}$ [32]

FRIDERICIA’s FORMULA: $QTcF = \frac{QT}{3\sqrt{RR}}$ [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 $<450$ msec, $451-470$ msec as borderline and $>470$ msec as prolonged. For men, $<430$ 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].](image)

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 Ikr 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 / Ikr 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 multicellular 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](image)

*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].](image)

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
Using a left ventricular wedge preparation of the canine heart, transmembrane 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).

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 $E_{\text{max}}$ 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 “pluridimensional 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.
### Table 1. Summary of current techniques and methodologies used for the assessment of cardiovascular safety

<table>
<thead>
<tr>
<th>Technology</th>
<th>Advantage</th>
<th>Limitation</th>
</tr>
</thead>
</table>
| **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 |
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].

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).
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].](image)

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


An Introduction to Mind the Gap


