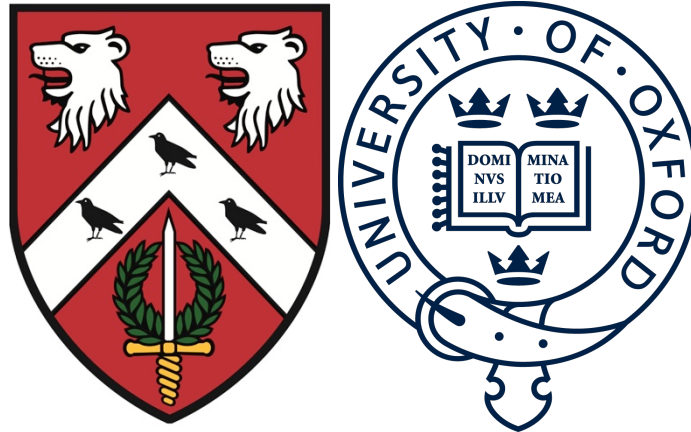


In silico methods for the prediction of drug-induced cardiotoxicity



Polina Mamoshina
St Anne's College

Computational Cardiovascular Science Group
Department of Computer Science
University of Oxford

Trinity Term 2020

This thesis is submitted to the Department of Computer Science, University of Oxford, for the degree of Doctor of Philosophy. This thesis is entirely my own work, and, except where otherwise indicated, describes my own research.

To my grandmother:

Посвящаю с любовью своей бабушке, Елене
Васильевне Беловой.

25.10.2019

Polina Mamoshina
St Anne's College

Doctor of Philosophy
Trinity Term 2020

In silico methods for the prediction of drug-induced cardiotoxicity

Abstract

Unexpected adverse reactions, especially unsafe cardiac effects, are a major concern of pharmaceutical companies that can prompt them to both discontinue drugs currently in development and withdraw drugs already on the market. Therefore, the safety assessment is a key stage of both the drug development process and the current regulatory framework of clinical trials. Given the importance of unforeseen acute electrophysiological effects in precipitating potentially lethal arrhythmias, the current preclinical testing stages of drug development are largely focused on their detection. However, a substantial number of drugs also affect cardiac function on many other levels, including contractility, mitochondria function and cell signalling. A number of in vitro, in vivo and in silico approaches capable of detecting different types of possible cardiovascular side effects have been proposed recently. Among those, human-based computational methods hold a great potential to increase the productivity of drug discovery pipelines, drive a more rational drug design and replace costly animal experiments that have limited translational ability for humans.

Therefore, the goal of this thesis is to propose a computational approach to predict drug-induced cardiotoxicity. A multi-label machine learning classification approach is used to simultaneously predict multiple forms of clinical cardiac side effects and take into account relationships between those forms of toxicity. In the last part of this thesis, the effects of trafficking

impairment, as one of the cardiotoxicity mechanisms, are then investigated using simulations of action potential models.

Publications

Below is a list of publications resulting from the work described in this thesis:

- “Dual transcriptomic and molecular machine learning predicts all major clinical forms of drug cardiotoxicity” (submitted). P. Mamoshina, A. Bueno-Orovio, B. Rodriguez.
- “Cardiotoxic drugs: classification of adverse reactions and mechanisms of action” (submitted). P. Mamoshina, A. Bueno-Orovio, B. Rodriguez.
- “Modelling of genetic mutations and drug action causing deficiency in hERG trafficking” (in preparation). P. Mamoshina, B. Rodriguez, A. Bueno-Orovio.

Conferences

The work described in this thesis was presented at the following conferences:

Oral presentations

- September 2017: “Signalling pathway analysis of drug-induced cardiotoxicity”. In Silico Drug Safety and Efficacy Symposium (Oxford, UK).
- November 2017: “Beyond electrophysiology: predicting drug-induced cardiotoxicity by combining transcriptome analysis and machine learning”. The Cardiac Physiome Society’s annual meeting (Toronto, Canada).
- November 2018: “Predicting drug-induced cardiotoxicity through machine learning – AI vs cardiotoxicity”. BioData world (Basel, Switzerland).
- December 2018: “Machine Learning for biomarkers of drug responses”. NeurIPS Expo Workshop series (Montreal, Canada).
- September 2019: “Artificial intelligence for safety pharmacology”. Safety Pharmacology Society Annual Meeting (Barcelona, Spain).
- February 2020: “Modelling of genetic mutations and drug action causing deficiency in hERG trafficking”. 6th UC Davis Cardiovascular Symposium (Sacramento, USA).

Posters

- November 2017: “Beyond electrophysiology: predicting drug-induced cardiotoxicity by combining transcriptome analysis and machine learning”. The Cardiac Physiome Society’s annual meeting (Toronto, Canada).
- March 2018: “Applications of machine learning in drug discovery”. Artificial Intelligence @ Oxford expo (Oxford, UK).
- March 2018: “Beyond electrophysiology: predicting drug-induced cardiotoxicity by combining transcriptome analysis and machine learning”. 5th Annual Oxbridge Women in Computer Science conference (Cambridge, UK).
- July 2018: “Prediction of all form drug-induced cardiotoxicity by combining transcriptome analysis, structural data and machine learning”. 18th World Congress of Basic and Clinical Pharmacology (Kyoto, Japan).
- September 2019: “Machine learning for six forms of cardiotoxicity prediction”. Safety Pharmacology Society Annual Meeting (Barcelona, Spain).

R package and database description

Dedicated R package ‘MLC’ along with curated database is available upon request. ‘MLC’ package can be used to fit multi-label models specified by a chain of predictors or a set of individual predictors. The model supports binary and multi-class multi-label classifications and multi-label regression. The included in the package database contains information about 1131 drugs summarized including name and identifiers (Broad Institute id, PubChem chemical id, DrugBank id), 7 molecular descriptors and fingerprint of molecules and transcriptional profiles where available. The detailed description of the data is provided in Section 4.2 (p. 82) of this thesis.

Contents

1	Introduction	13
1.1	Thesis goals	14
1.2	Outline of thesis	16
2	Introduction to drug-induced cardiotoxicity	19
2.1	Main molecular mechanisms of drug cardiotoxicity	20
2.1.1	Electrophysiology	22
2.1.1.1	Direct block	22
2.1.1.2	hERG trafficking inhibition	23
2.1.2	Contractility	25
2.1.3	Mitochondrial toxicity	26
2.1.4	Growth factors and cytokines signalling	28
2.2	Main classes of drugs causing clinical cardiotoxicity	29
2.2.1	Antineoplastic agents	29
2.2.1.1	Antimitotic	30
2.2.1.1.1	Anthracyclines	30
2.2.1.1.2	Cross-linking agents	30
2.2.1.1.3	Anti-tubulins	31
2.2.1.2	Tyrosine kinase inhibitors	31
2.2.1.3	Monoclonal antibodies	32
2.2.1.4	HDAC inhibitors	33
2.2.1.5	Other antineoplastic agents	33
2.2.2	Central nervous system agents	34
2.2.2.1	Local anesthetics	35
2.2.2.2	Antidepressants	35
2.2.2.3	Antipsychotic	36
2.2.2.4	Neurodegenerative disease agents	36
2.2.2.5	Other central nervous system agents	37
2.2.3	Genitourinary system agents	37
2.2.4	Gastrointestinal agents	38
2.2.5	Antihistamines	39
2.2.6	Anti-inflammatory agents	39
2.2.7	Anti-infective agents	40
2.2.7.1	Antibiotics	40
2.2.7.2	Antivirals	41

2.2.7.3 Other anti-infective	41
2.2.8 Cardiovascular agents	42
2.2.9 Other agents	43
2.3 Conclusions	43
3 Computational approaches for prediction and investigation of drug-induced cardiac side effects	64
3.1 Introduction	64
3.2 Gene expression signatures of drug-induced cardiac damages	64
3.3 Gene expression analysis	66
3.3.1 Gene list annotation	67
3.3.2 Gene list functional annotation	68
3.4 Applications of machine learning in cardiotoxicity prediction	70
3.4.1 Case studies	71
3.4.2 Challenges of applications of machine learning in biomedicine	73
3.5 Cardiac human cellular electrophysiological models	74
3.6 Cardiac cellular electrophysiological models of drug-induced effects	76
3.6.1 Biomarkers for quantifying cardiac electrophysiological model output	77
3.7 Databases for drug discovery	78
3.7.1 Pathway databases	78
3.7.2 Gene expression databases	79
3.7.3 Side effect databases	79
3.8 Conclusions	80
4 Prevalence of cardiotoxicity forms: data analysis	82
4.1 Introduction	82
4.2 Methods and Materials	82
4.2.1 Data collection	82
4.2.2 Calculation of molecular descriptors and fingerprints	85
4.2.3 Training testing set design	85
4.2.4 Feature selection methods	86
4.2.5 Measure of association	86
4.2.6 Visualisation of hidden patterns in the data	87
4.3 Results	88
4.3.1 Frequency analysis of cardiotoxicity forms	88
4.3.2 Analysis of drug-induced cardiac disorders' association	89
4.3.3 Visualisation of the dataset structure	91

4.3.4	Prevalence of specific forms of cardiotoxicity associated with drug pharmacological classes	96
4.4	Discussion	97
4.5	Conclusions	98
5	Dual machine learning prediction of drug-induced cardiotoxicity	99
5.1	Introduction	99
5.2	Methods	100
5.2.1	Study design	100
5.2.2	Training, validation and testing set design	102
5.2.3	Multi-label balancing	103
5.2.4	Chain of classifiers with nested stacking	104
5.2.5	A set of individual classifiers	106
5.2.6	Optimisation schema	106
5.2.7	Model evaluation	106
5.3	Results	111
5.3.1	Chain of classifiers versus a set of individual predictors	111
5.3.2	Random vs leave-drug-out cross-validation	113
5.3.3	Dual vs individual feature sets	114
5.3.4	Predictive accuracy across drug classes and cardiotoxicity forms	116
5.4	Discussion	120
5.5	Conclusions	122
6	Computational modelling and simulation of hERG protein dynamics in human cardiomyocytes	124
6.1	Introduction	124
6.2	Materials and Methods	124
6.2.1	Characterisation of hERG trafficking stages and key regulatory proteins	124
6.2.2	Formulation of the hERG trafficking model	125
6.2.3	Implementation of the model	126
6.2.4	Calibration of the model	126
6.2.4.1	Characterisation of in vitro hERG trafficking experiments	126
6.2.4.2	Estimation of model parameters	128
6.2.5	Sensitivity analysis of hERG trafficking model	130
6.2.6	Partial correlation analysis	130
6.2.7	Simulation protocol of action potential models of human iPSC-derived cardiomyocytes coupled to hERG trafficking	131

6.2.8	Action potential biomarkers	131
6.3	Results	132
6.3.1	Estimated parameters of the hERG trafficking model	132
6.3.2	Sensitivity of the hERG trafficking model	134
6.3.3	Simulations of impaired hERG trafficking	136
6.4	Discussion	140
6.5	Conclusions	142
7	Simulations of the effects of drug-induced trafficking inhibition on cell electrophysiology	143
7.1	Introduction	143
7.2	Materials and Methods	144
7.2.1.1	Characterisation of gene expression profiling experiments	144
7.2.1.2	Characterisation of in vitro hERG trafficking experiments of drug incubation	144
7.2.2	Gene expression and pathway analysis	145
7.2.3	Calibration of the trafficking model	149
7.2.3.1	Estimation of trafficking model parameters	149
7.2.4	Validation of the trafficking model	150
7.2.4.1	Characterisation of action potential recordings of drug incubation experiments	150
7.2.4.2	Simulation protocol of action potential models of hiPSC-derived cardiomyocytes and coupling with hERG trafficking model	151
7.2.4.3	Control population of hERG trafficking models	152
7.2.4.4	Control population of human iPSC-derived cardiomyocytes models and coupling with hERG trafficking models	154
7.2.4.5	Action potential biomarkers of drug action	156
7.2.5	Detection of repolarisation abnormalities	156
7.2.6	Partial correlation analysis	157
7.2.7	L-type calcium current activation by arsenic trioxide	157
7.3	Results	158
7.3.1	Gene expression and pathway analysis of cell responses to incubation with arsenic trioxide	158
7.3.2	Simulation of arsenic trioxide action on hERG protein expression	160
7.3.3	Simulation of arsenic trioxide action on cellular electrophysiology	165
7.3.3.1	Arsenic trioxide effects on hERG trafficking under variability in hERG trafficking properties	170
7.3.3.2	Arsenic trioxide effects on hERG trafficking under variability in hERG trafficking and cellular electrophysiology properties	173

7.3.3.3	Arsenic trioxide effects on hERG trafficking and I _{CaL} current under variability in hERG trafficking and cellular electrophysiology properties	176
7.4	Discussion	177
7.5	Conclusions	179
8	Conclusions	180
8.1	Summary of the thesis	180
8.1.1	Key contributions and main findings	183
8.2	Future directions	184
8.2.1	Incorporating information about related cardiac side effects	184
8.2.2	Further experimental validation of the dual machine learning predictor of six cardiotoxicity forms	184
8.2.3	Simulation of hERG trafficking inhibition and acute hERG block	185
8.2.4	Simulation of trafficking-deficient mutations of hERG	186
8.2.5	Trafficking modelling of other channels	187
8.2.6	Tissue and organ effect of trafficking inhibition	188
8.3	Conclusions	188
9	Bibliography	190

Acknowledgements

I want to thank my two supervisors, Professor Blanca Rodriguez and Dr Alfonso Bueno-Orovio, who guided and supported me through this thesis. Blanca, thank you very much for all the support and helping me to focus on the important and enjoyable parts of the research. Alfonso, thank you for sharing your passion for science when I needed it most and your patience. This work would never have been done without your help. It has been an amazing three years, not without challenges, but I would not trade my time in Oxford for anything.

This work would be possible without the help and support of many collaborators. Special thanks to Dr Alex Zhavoronkov, who encouraged me to pursue the DPhil and all the support he provided. I would also like to thank Dr Michelangelo Paci for providing action potential model of Human iPSC Derived Cardiomyocytes and his guidance in simulations. Thank you to the Department of Computer Science for a fantastic workplace and Insilco Medicine for their support of my thesis and for providing computational resources.

I thank my Oxford friends and colleagues for vivid discussions we had. Cristian, your positivity kept me going and helped me to stay on a bright side so many times. Michelangelo, thank you for our rare but always excellently spent evenings and all the heart-to-heart talks we had. And thanks to everyone in Computational Cardiovascular Science Team for such an amazing work environment.

I would like to thank my family, especially my dad and mum for their support of my goals and dreams and their love in my life, my grandparents for their care and my brothers for their understanding.

Finally, I would like to thank my two little black cats, which are always in my heart.

1 Introduction

According to World Health Organization (WHO) statistics for the year 2015, cardiovascular diseases are one of the leading causes of death worldwide [1]. But despite the huge burden of cardiovascular diseases, the development of medications that could cure or prevent cardiac complications accounts for a small proportion of new medical entities compared to other disease areas, such as cancer [2,3]. Coupled with an overall high failure rate in clinical trials, the number of new drugs approved for cardiology and vascular diseases is disturbingly low [3], while the need for novel therapies remains high. Moreover, cardiovascular drugs, along with cancer drugs, have one of the lowest probabilities (lower than the median across indications) of being approved in clinical trials [4].

Lack of cardiac safety is among the top reasons for the discontinuation of clinical trials and the marketing of medications [5]. Any factors that increase the already high chances of developing cardiovascular complications would have a significant impact on the morbidity and mortality of patients. Therefore, there is a clear need for new approaches to improve the assessment of drug effects, both for efficacy and safety testing.

With advances in modern molecular biology, drug development has now extensively employed in vivo and in vitro testing of drug candidates to obtain information on metabolism, efficacy and potential toxicity. Although animal models still play an important role in drug discovery, the overall high cost and low translation rate raise the need for more humanised models [6,7].

Current open-source repositories of omics data already provide tens of thousands of high-quality samples for analysis in order to drive a more rational drug design [8]. Greater data availability coupled with algorithmic advances and a rise in computational power have led to an increase in the development of new computational methods, which now occupies an important place in drug discovery pipelines in both preclinical research and clinical trials. Machine learning methods are gaining popularity as powerful

tools to analyse biomedical data, including starting from target identification to prediction of clinical trial outcomes [9–11]. Similarly, modelling and simulation approaches ranging from single-cell to whole-heart models have been proven useful in verifying important insights into drug safety and efficacy, as well as patient stratification [12].

Therefore, the development of computational approaches for drug-induced cardiac-adverse reactions has a potential application in reducing animal experiments and in optimising the preclinical testing of drug safety.

1.1 Thesis goals

The main goal of this thesis is to develop computational methods to predict drug-induced cardiotoxicity. To address this goal, the thesis considers two different but complementary methodologies: machine learning predictors and human cardiomyocyte modelling and simulation. Machine learning predictors allow to generate models from data without prior knowledge of biological dependencies to be determined, without the assumption that a single unique model exists. The second approach builds on simulations using physiologically-relevant human cardiomyocyte cellular models to identify mechanisms underlying cardiac side effects.

To achieve its overall goal, the thesis addresses three specific research questions:

1. Can we predict multiple drug-induced cardiac disorders simultaneously using transcriptional profiles and structural information of drugs?
2. Can we model hERG trafficking and explore how alterations of its stages would affect cellular electrophysiology?
3. Can we predict the electrophysiological effects of drugs inhibiting trafficking of hERG on human induced pluripotent cardiomyocytes?

The first part of this thesis was to collect a large comprehensive dataset of compounds with transcriptional profiles, derived molecular descriptors and fingerprints, as well as information about their cardiac safety in human. With this dataset, an independently validated, supervised multi-label machine learning was developed. The algorithms included the drug characteristics as input and identified multiple forms of cardiotoxicity as output. I initially collected a large dataset of 1,131 drugs, with reports on their clinical cardiac safety and the presence of adverse cardiac reactions. The primary analysis of this dataset demonstrates the association between different cardiotoxicity types induced by drugs and that compounds with cardiac complication reports are not linearly separable. These findings then motivated the development of a machine learning approach capable of capturing the dependencies between drug characteristics and their cardiac safety profile by incorporating information about the relationship between cardiotoxicity types, as described in Chapter 5. The proposed approach demonstrated good predictive accuracy on previously unseen drugs and new cell lines.

Whereas the machine learning algorithm is able to identify drug-induced cardiotoxic effects, it is unable to explain the underlying mechanisms. Thus, a second approach is followed in the second part of the thesis, to investigate mechanisms underlying a specific cardiotoxicity type. Here I focus on the inhibition of hERG trafficking and its pro-arrhythmic potential, using simulations combining a novel model of hERG protein dynamics and physiologically-based cardiac cellular electrophysiology models. A new model of hERG protein trafficking is developed to simulate the dynamics of mature functional hERG on the cellular membrane as a function of time. This allows to compare normal versus altered trafficking by drugs, or other pathological conditions, based on their mechanisms of action derived from their gene and protein expression profiles. By inserting the amount of mature functional hERG into a population of action potential models, the interactions between altered trafficking, direct block and other electrophysiological components of cardiomyocytes are studied, and the pro-arrhythmic potential of the drugs tested.

1.2 Outline of thesis

In **Chapter 2**, I present the background for the thesis by introducing an overview on drug-induced cardiotoxicity. I briefly cover all main mechanisms of cardiotoxicity, including the disruption of electrophysiology, contractility, mitochondrial toxicity and growth factors and cytokines signalling. I move on to describing the pharmacological classes of 77 drugs, covering side effects and mechanisms of anticancer, central nervous system, genitourinary, gastrointestinal, antihistamines, anti-inflammatory, and anti-infective agents. I finish the review by highlighting their common modes of action on cardiac function.

Chapter 3 provides an introduction to the computational analysis of drug-induced side effects. I start with a literature review on the gene expression analysis of transcriptional signatures of cardiotoxic drugs. After this, I provide an introduction to signalling pathway analysis. I then summarise case studies describing the application of machine learning in stratifying safe drugs from drugs that are capable of causing clinical complications. Following this, I describe the key challenges in building machine learning predictors. In particular, I cover the differences between cross-validation strategies. Following this, I review current approaches for the mechanistic modelling of drug-induced adverse reactions with cardiac cellular electrophysiological models and describe the set of key biomarkers used to characterise the outputs of those models. Finally, I describe the existing databases for computational drug discovery.

Chapter 4 presents the first results in this thesis and the research done in data collection and primary analysis. I describe the curation of the dataset from the public domain, as well as databases and knowledge sources such as MedDRA, MESH, Connectivity map, SIDER and Drug Bank. The collected dataset contains transcriptional profiles, structural information and protein target information about marketed drugs reported to have clinical cardiac-adverse reactions, and investigational and withdrawn drugs from the market. Afterwards, I describe the analysis of the dataset. In this study, I explore the

co-occurrence of multiple types of cardiac side effects and their prevalence in specific drug classes. I show that there is an association between different forms of adverse drug-induced cardiac effects as a motivation for the model proposed in Chapter 5.

In **Chapter 5**, I exploit the database collected in Chapter 4 and I present a multi-label predictor of drug-induced cardiotoxicity that uses both structural information (molecular descriptors and fingerprints) and transcriptional profiles of drugs. The proposed approach employs a model of chain of classifiers with nested stacking, which accounts for association between cardiotoxicity forms. I compare this approach with a set of individual classifiers to confirm that it is important to preserve the existing associations between cardiotoxicity forms for better discrimination between safe drugs and drugs with reported cardiac-adverse reactions. In this chapter, I also analyse the predictive performance of the proposed approach across different pharmacological classes of drugs and cell lines.

In **Chapter 6**, I move from a machine learning-based approach to mechanistic modelling, and propose a physiological model of hERG trafficking in cells. I start with a description of the key stages of trafficking. After this, I move to the formulation of the model and its description as a set of ordinary differential equations. I then introduce the calibration of the model with experimental data and quantification of model parameters. I explore the sensitivity of the model, and by coupling it with a cardiac electrophysiological model, I use it as a method for hypothesis generation on how changes in its parameters would affect the levels of hERG on the cellular membrane, and would therefore manifest changes in action potential in the case of pathological conditions (trafficking mutations) or drug-induced trafficking inhibition.

In **Chapter 7**, I provide simulation results on the effect of the inhibition of hERG trafficking by arsenic trioxide, as an exemplar of its application for investigations on pharmacological action. I first use gene expression to identify the mechanism of the drug action in hERG impairment, and protein data to estimate its trafficking inhibition rate at the protein level. I then use this

information to simulate arsenic trioxide effects on a single-cell level and on a population of experimentally calibrated models of cardiac action potential models of human iPSC-derived cardiomyocytes. I compare the simulated changes on action potential to experimental electrophysiological recordings of human iPSC-derived cardiomyocytes incubated with arsenic trioxide for 24 hours, and find a good agreement. In addition to hERG trafficking inhibition, I also investigate upregulation of the L-type calcium channel, shown for arsenic trioxide in animal experiments.

In **Chapter 8**, I conclude the thesis with an overview of its major findings and contributions. I then finish this thesis with a discussion of the limitations of the current research methodology and obtained results, and provide suggestions for potential future studies that were beyond the scope of this thesis.

2 Introduction to drug-induced cardiotoxicity

Concerns regarding cardiac safety are among the top reasons for the withdrawal of drugs from clinical trials and the market. Despite extensive efforts to elucidate the links between adverse cardiovascular events and pharmaceutical agents, cardiotoxicity and its constituent molecular processes have yet to be comprehensively described. The severity of cardiac toxicity varies from myocardium dysfunction to terminal heart failure. In addition, its effect varies among patients, and cardiotoxicity has been shown to depend on the gender, age [13,14], and genetic background of the individual [15–18].

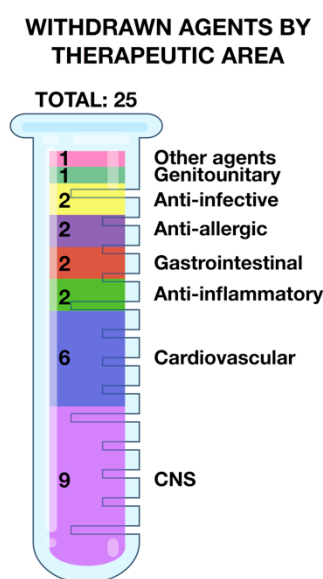


Figure 2-1 Number of drugs withdrawn from market due to cardiotoxicity up to January 2020, according to their therapeutic areas.

Drug-induced cardiotoxicity imposes a substantial limit on the application of a myriad of drugs. Although most cardiotoxic events are due to unforeseen electrophysiological activity [19], recent studies have indicated that a considerable number of drugs disrupt cardiac function by impairing myocardial metabolism and cardiac structure. In this chapter, I summarise the drug-protein mechanisms by which drug-induced cardiotoxicity may develop. I provide the review of 77 cardiotoxic drugs, 25 of which have been withdrawn from the market due to cardiotoxicity (Figure 2-1 and Table 2-1), covering side effects and cardiotoxic mechanisms of anticancer, central nervous system, genitourinary, gastrointestinal, antihistamines, anti-inflammatory, and anti-infective agents, furthermore demonstrating they have common modes of action on cardiac function and on adverse cardiac events.

2.1 Main molecular mechanisms of drug cardiotoxicity

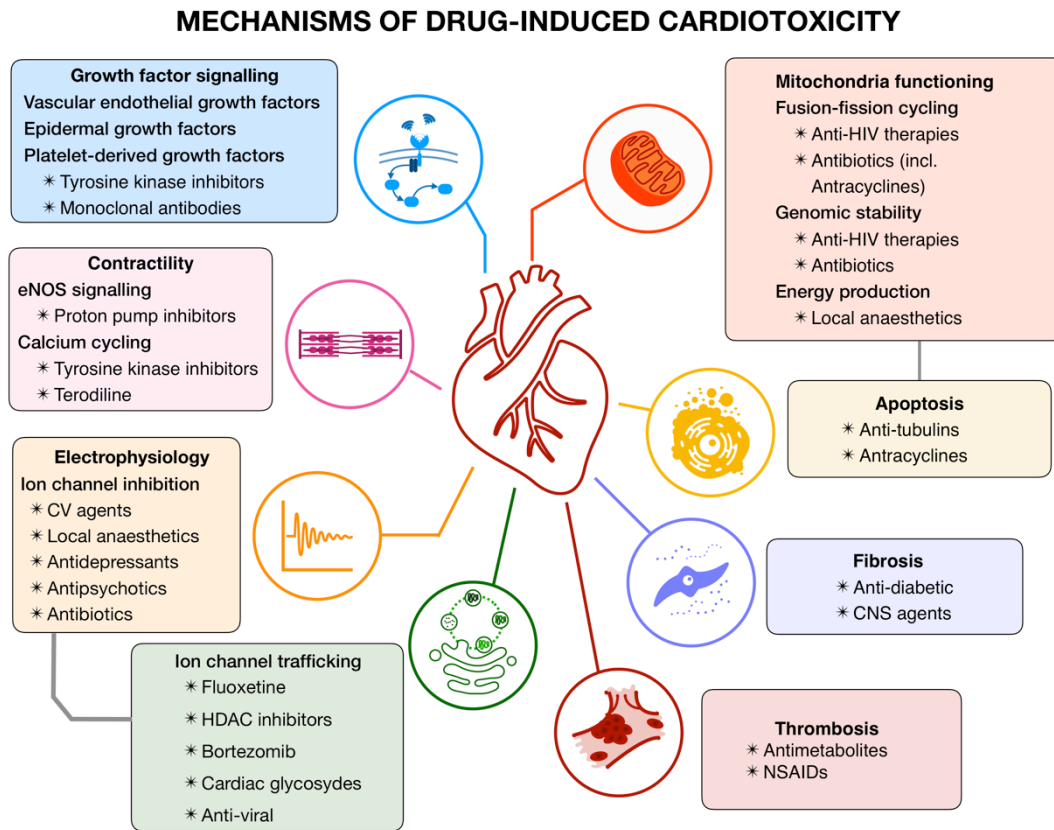


Figure 2-2 Mechanisms of drug-induced cardiotoxicity. Antibiotics (including anthracyclines) and antiviral therapies induce mitochondrial dysfunction, leading to mitochondrial fragmentation and an impaired fusion-fission cycle. Local anaesthetics interact with mitochondrial membranes and increase their permeability, affecting function. Cardiotoxicity of tyrosine kinase inhibitors and monoclonal antibodies is primarily linked to their inhibition of major signalling pathways responsible for cardiomyocyte survival and maintenance. Administration of a number of CNS and anti-diabetic agents has been associated with the emergence of cardiac fibrosis. The mechanism of proton pump inhibitors affecting cardiac contractility has been linked to the inhibition of NO synthesis. Similarly, inhibitors of tyrosine kinases have been linked to dysregulation of calcium cycling. Multiple neural agents, including selective serotonin reuptake inhibitors and antipsychotic drugs, cardiovascular agents, and antibiotics have been shown to affect cardiac tissue interacting with one or several cardiac ionic channels. At the same time, cardiac side effects of fluoxetine, HDAC inhibitors and bortezomib are primarily linked to their effects on cellular hERG trafficking rather than acute channel block. CNS: central nervous system; CV: cardiovascular; NSAIDs: non-steroidal anti-inflammatory drugs; HDAC: histone deacetylase; eNOS: endothelial nitric oxide synthase.

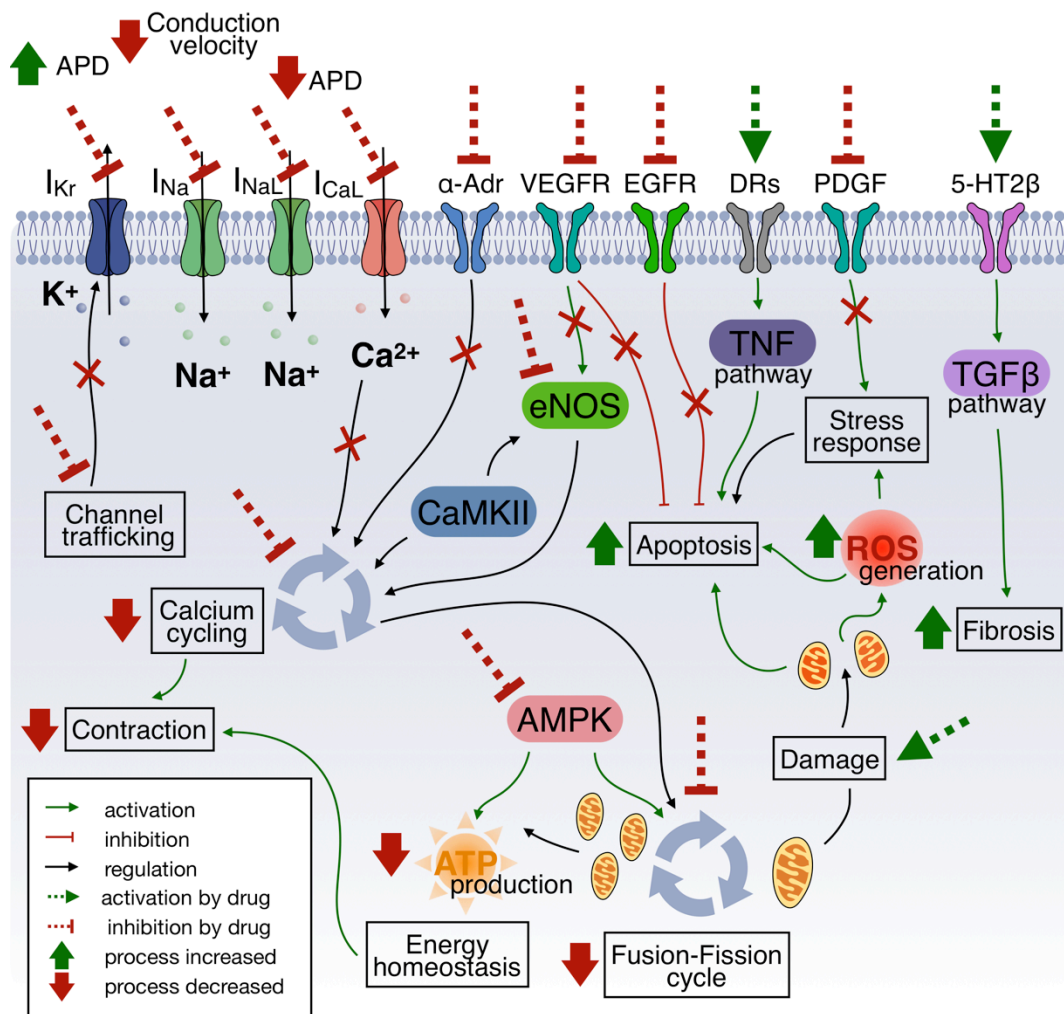


Figure 2-3 General overview of mechanisms of cardiotoxicity. Cardiotoxic compounds can exert their influence on cardiomyocytes through a variety of mechanisms. Abnormalities in action potential duration (APD) or conduction velocity are associated with a direct block of ionic currents (IKr, INa, ICaL, among others) or inhibition of their trafficking from the nucleus to the cell membrane. Apoptotic events can be triggered via multiple pathways, including inhibition of VEGFR or EGFR (which promote cardiomyocyte survival), PDGF inhibition (responsible for the compensatory stress response) or activation of DR-induced TNF signalling. Mitochondrial damage (caused by suppression of the fusion-fission cycle or effects on mitochondrial DNA) and elevated ROS levels also initiate apoptosis. Inhibition of α -Adr, VEGFR signalling, or direct eNOS inhibition affect calcium cycling. Inhibition of AMPK signalling has effects on both the mitochondrial fusion-fission cycle and ATP production. Serotonin-induced activation of TGF β pathway via the 5-HT2 β receptor is primarily linked to the induction of cardiac fibrosis. APD: action potential duration; IKr: rapid delayed rectifier current; INa: inward sodium current; INaL: late inward sodium current; ICaL: L-type calcium current; α -Adr: alpha-adrenergic receptor; VEGFR: vascular endothelial growth factor receptor; EGFR: epidermal growth factor receptor; DRs: death receptors; PDGF: platelet-

derived growth factor; 5-HT₂ β : 5-Hydroxytryptamine receptor (serotonin receptor 2B); TNF: tumor necrosis factor; TGF β : transforming growth factor beta; eNOS: endothelial nitric oxide synthase; ROS: reactive oxygen species; CaMKII: calcium/calmodulin-dependent protein kinase II; AMPK: AMP-activated protein kinase; ATP: adenosine triphosphate.

Cardiotoxicity generally results from the simultaneous interruption of key myocardial functions and viabilities. Here I draw focus to unexpected drug-induced effects on the myocardium, via the disruption of electrophysiology, contractility, mitochondrial toxicity, growth factor, and cytokine regulation (Figure 2-2), through complex cell signalling pathways (Figure 2-3).

2.1.1 Electrophysiology

2.1.1.1 *Direct block*

In the last decades, preclinical testing for cardiac adverse events has focused on screening assays of hERG channel inhibition [20], which are aimed to detect prolongation of the action potential. However, the biological role of hERG channels is not only limited to their mediation in cardiac repolarisation via the 'rapid' delayed rectifier current (I_{Kr}). hERG encoded channels are also involved in cell proliferation and malignant cell apoptosis [21]. For instance, the expression of I_{Kr} in gliomas has been recently proposed as a specific target for antineoplastic therapy [22]. However, drugs often interact with multiple channels at once, altering electrophysiology in a multifactorial way [23–25]. Recent work regarding the effect of drugs on multiple ion channels and connections to adverse reactions has resulted in significantly improved toxicity prediction, paving the way for safer pharmaceutical treatment. There is evidence to suggest that the electrophysiological implications of drug-induced hERG block could be alleviated by drug interactions with other ionic channels [26,27]. Nevertheless, the absence of hERG block does not necessarily imply that a molecule is incapable of producing adverse cardiac effects [28,29]. Both the amplitude and density of cardiac currents depend on, among other factors, cell type, age, and cardiac health. Pathophysiological conditions such as diabetes have been shown to abate the amplitude of a number of ion channel currents [30], potentially enhancing susceptibility to adverse reactions. Multiple neural agents, including selective serotonin reuptake

inhibitors and antipsychotic drugs, have been also shown to simultaneously interact with several cardiac ionic channels.

2.1.1.2 *hERG trafficking inhibition*

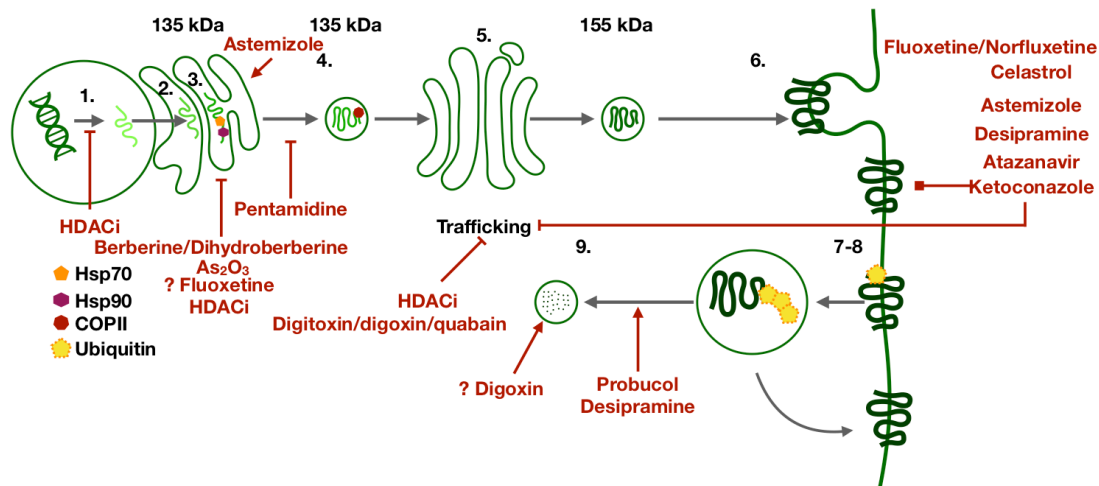


Figure 2-4 Schematic of hERG biogenesis, trafficking and degradation, and pathways of drug-induced IKr deficiency. Subcellular processes modulating hERG expression on the cellular membrane include: 1) mRNA synthesis, 2) synthesis of polypeptide chain, 3) transport of polypeptide from endoplasmic reticulum to Golgi apparatus with Hsp70 and Hsp90 chaperones attached, 4) transport of the polypeptide in the COPII-mediated vesicle, 5) glycosylation, 6) exocytosis, 7) endocytosis, 8) ubiquitination, and 9) degradation of hERG.

In addition to direct block of the hERG channel, a number of studies have shown that multiple pharmacological agents can cause hERG deficiency by also (or only by) the inhibition of its biogenesis and trafficking [31,32], as shown in Figure 2-4 and Table 2-1. Compared to the fast action of acute drug effects, hERG trafficking impairment manifests as long-term effects, with time scales of hours to days.

For example, arsenic trioxide (Figure 2-4), extensively used in cancer treatment, interacts with chaperones Hsp70 and Hsp90 altering the folding process of hERG. As a result, it has been shown to affect the maturation of the channel in human embryonic kidney 293 cells with wild-type hERG (hERG-HEK293) [33]. Folding inhibition also has been proposed as a potential mechanism of the dual tumour suppression and carcinogenicity of arsenic trioxide [34]. Berberine, an antidiarrheal medicine and its derivative,

dihydroberberine (Figure 2-4), have been shown to inhibit hERG trafficking by impairing folding in both hERG-HEK293 cell lines and in guinea pig ventricular cardiomyocytes [35,36]. The inhibition of folding by those drugs results in the reduction of the mature glycosylated form of 155 kDa hERG, which can be easily measured along with immature hERG (135 kDa). Accumulation of unfolded proteins activates the unfolded protein response pathway, which includes activating the transcriptional factor 6 (ATF6) pathway, as shown in hERG-HEK293 cells incubated with berberine [35] and in mouse vascular endothelium cell lines incubated with arsenic trioxide [37]. For the anti-infective drug pentamidine (Figure 2-4), the reduction in mature hERG form in hERG-HEK293 cell lines was identified to be a result of the suppression of transport of the immature form, rather than a folding alteration [38]. The same mature hERG deficiency was observed in that study in neonatal rat ventricular cardiomyocytes incubated with pentamidine.

Cardiac glycosides (Figure 2-4), specifically Na^+/K^+ -ATPase inhibitors with known QT prolongation risk, reduce hERG trafficking in hERG-HEK293 cell lines and in guinea pig ventricular cardiomyocytes [39]. Incubation with digitoxin, digoxin and ouabain resulted in a significant reduction of the mature channel form. Notably, Kv1.5 and IKs trafficking were not affected by those glycosides, suggesting that a complex mechanism is involved. Remarkably, the hERG blocker astemizole is able to rescue digitoxin, digoxin and ouabain-induced hERG deficiency, possibly acting as a chaperone [40]. Later, evidence was provided that digoxin-induced hERG trafficking inhibition is a result of the reduced potassium concentration caused by Na^+/K^+ -ATPase inhibition [41]. Moreover, digoxin has been reported to enhance protein degradation in general through a lysosomal pathway [42].

A lipid-lowering compound, probucol, is also known to cause QT prolongation without a direct block of IKr (Figure 2-4). This effect is believed to be a result of trafficking disruption, as shown in neonatal rat ventricular cardiomyocytes [43] and later in hERG-HEK293 cells, where enhanced degradation of the mature channel via neddylation was indicated as a key mechanism [44]. Meanwhile, arsenic trioxide showed no effect on neddylation of hERG [44].

Gene expression inhibition and alteration of the acetylation levels of a set of proteins involved in ion channel trafficking and degradation have been also proposed as possible cardiotoxic mechanisms of HDAC inhibitors [45].

In addition to the aforementioned mechanisms, a substantial number of drugs are known to bind to the IKr channel, inhibiting it and disrupting the trafficking of the mature channel at the same time. One such example is ketoconazole (Figure 2-4), an anti-infective agent that causes clinical QT prolongation, shown to co-inhibit IKr and hERG trafficking in hERG-HEK293 cells [46]. Interestingly, a lack of mature hERG was observed in both wild-type and mutated channels that do not bind to the drug, suggesting trafficking impairment as a main mechanism. Such mechanism was however not studied in detail, and the biological assays showed inconsistent results with inhibition of the Hsp90 chaperone in yeast strains (Pubchem: AID 540270), but not in the cancer cell lysate Hsp90 (Pubchem AID 429) or rabbit cell lysate Hsp90 (Pubchem AID 1789). A similar effect in hERG-HEK293 cells was shown by atazanavir, another anti-infective agent with known risk of causing Torsade de Pointes arrhythmias (Figure 2-1; [47]), and by the antidepressants fluoxetine and norfluoxetine (Figure 2-1; [49]). Interestingly, fluoxetine can induce elevation of Hsp90 in mouse brain endothelial cells [50], but no previous research has investigated such effects either in cardiac or even in hERG-HEK293 cells. In contrast, the mechanisms of the antidepressant desipramine (Figure 2-4), with dual action on activity and trafficking, have been well characterised [51]. Similarly to probucol, desipramine accelerates degradation of mature hERG and does not interfere with Hsp90 and Hsp70 chaperones in hERG-HEK293 cells. A study of celastrol effects on HEK293 cells expressing Kir2.1, hERG and Kv2.1 channels showed that the compound inhibits both the function and trafficking of Kir2.1 and hERG, but not Kv2.1 [52]. Interestingly, the effect on trafficking was more potent (5-10 fold) than the acute block of the channel.

2.1.2 Contractility

Contractility, the most essential heart function, has been observed to be disrupted by numerous pathophysiological conditions, including drug-induced cardiotoxicity. Cardiomyocytes connect electromechanically with each other and the extracellular matrix via gap junctions, which are key molecular regulators for coordinated contraction and intercellular communication. In addition, homeostatic calcium cycling in cardiomyocytes is essential for normal cardiac function [53]. Calcium cycling regulation is a complex process involving multiple regulatory molecules. Specifically, calcium/calmodulin-dependent protein kinase II (CaMKII) regulates calcium cycling via phosphorylation of its targets on both the cellular membrane and the sarcoplasmic reticulum. Accordingly, CaMKII inhibition has been shown to have potential efficacy for antiarrhythmic therapy [54].

An additional molecular signal for cardiac muscle contraction, Nitric Oxide (NO), further regulates intracellular Ca^{2+} concentration, promotes vascular relaxation, and inhibits platelet aggregation. In cardiomyocytes, NO is primarily synthesised by endothelial Nitric Oxide Synthase (eNOS), whose activity is regulated by intracellular CaMKII levels. Protein phosphatases, which act in opposition to CaMKII, are believed to regulate cardiac gap junctional communication [55]. When functioning normally, these processes coordinate to establish a homeostatic system, the destabilisation of which underlies one of the proposed mechanisms for the cardiotoxicity of proton pump inhibitors. It has been suggested that proton pump inhibitors inhibit dimethylarginine dimethylaminohydrolase (DDAH), an enzyme responsible for eliminating asymmetric dimethylarginine (ADMA), resulting in excess of ADMA impeding NO synthesis [56]. Similarly, antineoplastic agents such as tyrosine kinases have been linked to upregulated CaMKII expression and activity [57].

2.1.3 Mitochondrial toxicity

Representing approximately one third of cardiomyocytes by volume, mitochondria play an important role in cardiac function, mainly by satisfying

the immense energy requirements of cardiac contraction. As such, failure to replace malfunctioning mitochondria (mitochondrial toxicity) is highly injurious to cardiac function. To this end, altered cardiac metabolism has been linked to the development of cardiac diseases, such as ischemia [58], and conditions that increase the risk thereof, such as diabetes [59].

As a by-product of their metabolic function, mitochondria produce reactive oxygen species (ROS), which also play an important role in pro-apoptotic signalling [60]. These dynamic organelles engage in the mitochondrial fusion-fission cycle, changing in shape and number within the cell to respond to environmental stressors. Whereas mitochondrial fusion and fission help to reduce cellular stress in response to mild environmental stressors, these processes can also result in apoptosis in response to extreme stressors [61]. When exposed to disease conditions, increased mitochondrial fission (fragmentation) leads to tissue necrosis. Demonstrative of their tight association, the contractility of cardiomyocytes and the mitochondrial fusion-fission system were recently found to be closely coupled: the fusion-fission cycle depends on calcium oscillations, and abnormalities in the fusion-fission system lead to aberrant contraction [62,63].

In recent decades, several pharmaceutical agents have been found to produce or facilitate mitochondrial toxicity in various organ and tissue systems, including the heart [64]. These inimical drugs span multiple classes, including anthracyclines, antivirals, antidepressants, and local anaesthetics. The mitochondrial toxicity of anthracyclines is linked primarily to the inhibition of their direct target, topoisomerase (DNA) II Beta (TOP2B), which is required for mitochondrial DNA replication [65]. Antiretroviral drugs, often used to treat HIV, are designed to directly target and inhibit reverse transcriptase. However, antiretrovirals have been found to also inhibit mitochondrial DNA polymerase gamma (DNA Pol- γ), leading to depleted mitochondrial DNA and disrupted mitochondrial function [66]. Moreover, anti-HIV therapies have also been shown to inhibit quality control systems responsible for regulating the mitochondrial fusion-fission cycle [67]. In the case of local anaesthetics, mitochondrial toxicity has been suggested to occur due to their interactions

with phospholipids on the mitochondrial membrane, which often result in increased membrane permeability, electron transport chain disruption, and calcium accumulation [68,69].

2.1.4 Growth factors and cytokines signalling

Growth factors and cytokines are biologically active molecules that act directly on many cellular functions such as adhesion, proliferation, and migration. As a result, growth factor and cytokine signalling broadly affect tissue and organ function.

Vascular endothelial growth factor (VEGF) is a signal protein responsible, primarily via the eNOS signalling pathway, for the regulation of vascular formation, angiogenesis, cardiomyocyte development and proliferation, and myocardial regeneration [70]. In addition, VEGF is a major anti-apoptotic factor that promotes cardiomyocyte survival in response to environmental stress or disease [71,72]. Since angiogenesis has been observed to be closely related to neoplastic metastasis and malignancy, controlling angiogenesis via inhibition of VEGF signalling is an attractive target [73]. VEGF-inhibitory therapies, per contra, have been shown to produce many adverse side-effects, including cardiotoxicity [74]. In addition to the VEGF molecule itself, VEGF signalling could be triggered by immunomodulating factors, such as histamines, acting on the H1 receptor. Therefore, antihistamines have been proposed to exhibit angiogenic-inhibitory properties [75,76], but their cardiotoxicity linked to VEGF inhibition has yet to be studied.

The epidermal growth factor receptor (EGFR) is a receptor belonging to the EGF family of protein kinases, whose members include HER2, HER3, and HER4, and is essential to almost every known cellular process. Among these four proteins, HER2 is particularly crucial for cardiomyocyte differentiation and embryonic cardiac development [77]. EGFR signalling is antiapoptotic, indicating that it could lead to uncontrolled growth or intensified oncogenesis. To this end, there is sufficient evidence to suggest that EGF-activated VEGF

signalling may enable regulation of both angiogenesis and eNOS signalling [78]. Taking these factors together, it follows that targeting EGFR to inhibit neoplastic growth has become an attractive avenue for potential antineoplastic therapies [79]. Unfortunately, EGF inhibitors, such as lapatinib or trastuzumab, have recently been linked to cardiac adverse reactions [80]. Closely related to the VEGF, the platelet-derived growth factor (PDGF) is another key factor for proper myocardial development [81]. Research using mouse models has demonstrated that the PDGF also plays a role in the angiogenic cardiovascular compensatory stress response [82] and in cardiac fibrosis [83]. Inhibition of PDGF by sunitinib has been linked to the cardiotoxicity of this drug [74].

AMP-activated protein kinase (AMPK) is an enzyme that controls cellular energy and survival homeostasis. In addition, the active form of AMPK is required for mitochondrial fission in response to stress [84]. AMPK activation and the subsequent induction of compensatory mechanisms could therefore underlie the observed relatively low toxicity rate of lapatinib [80], compared to the disruption of myocardial metabolism associated to attenuation of AMPK signalling by other tyrosine kinase inhibitors [74].

2.2 Main classes of drugs causing clinical cardiotoxicity

2.2.1 Antineoplastic agents

The administration of cytotoxic drugs is always coupled with the risk of developing different adverse reactions. While significant improvements have been made in recent years regarding the overall safety of chemotherapy, cardiovascular damage remains as one of the most common side-effects of antineoplastic agents, and often results in diminished life expectancy for cancer patients [85]. Two types of chemotherapy-induced cardiotoxicity have been established: irreversible, referred to as Type I and characterised by cellular damage, and reversible, referred as Type II and characterised by cellular dysfunction [86].

2.2.1.1 Antimitotic

2.2.1.1.1 Anthracyclines

Among chemotherapeutic agents, anthracyclines have been shown to be effective for the widest range of antineoplastic applications. Anthracyclines have four mechanisms of action by which they target actively proliferating neoplastic cells: DNA-RNA intercalation, topoisomerase 2 (TOP2B) inhibition, iron-mediated generation of free radicals, and inducing the eviction of histone from chromatin [87].

The cumulative and dose-related cardiotoxicity of anthracyclines is primarily linked to ROS formation, mitochondrial dysfunction, and cardiomyocyte apoptosis [88,89]. Doxorubicin has been additionally shown to upregulate the expression of tumor necrosis factor (TNF) mediated death receptors, involved in mediating apoptosis in induced pluripotent stem cell-derived cardiomyocytes [90]. While cumulative cardiotoxicity is common among this class of compounds, there are several anthracyclines claiming to show no evidence of anthracycline-induced toxicity. This is for example the case of the FDA approved amrubicin, a drug marketed as “Calsed” in 2011 for lung cancer treatment.

2.2.1.1.2 Cross-linking agents

Antimetabolites, among the first effective chemotherapeutic agents, work by suppressing the division of potentially cancerous, quickly dividing cells via interference with DNA replication. Due to their mechanism of action, the most frequent side-effects of antimetabolites are observed in highly proliferative tissues such as those found in the gastrointestinal tract, skin or hair. Although less common, if cardiotoxicity arises from antimetabolite administration, it generally does within a week of therapy with severity ranging from symptomatic arrhythmias to sudden cardiac death [91,92]. Additionally, antimetabolites affect the coagulation system; 5-fluorouracil (5-FU) and cisplatin activate thromboxane A₂ (TXA₂) formation and platelet aggregation, leading to cardiac remodelling and ischemia [93,94]. 5-FU and cisplatin are

also known to cause electrolyte imbalances, which play an important role in gastrointestinal toxicity [95], and are believed to be involved in their cardiotoxicity [96]. Cytarabine is another antimetabolic agent with known cardiotoxic effects [97,98]; however, an understanding of cytarabine-induced cardiotoxic mechanisms is still lacking.

2.2.1.1.3 Anti-tubulins

Vinca alkaloids and taxanes, well-established clinical antineoplastic drugs, work to eliminate tumours by binding to microtubules. Similar to antimetabolites, *vinca* alkaloids and taxanes target cell division and are linked primarily to gastrointestinal toxicity. In addition, the administration of vinblastine has been associated with myocardial ischemia and infarction [99]. Similarly, paclitaxel and docetaxel have been associated with bradycardia, ischemia, and heart failure [100]. One of their proposed mechanisms of cardiotoxicity involves the inhibition of actively-proliferating cardiac endothelial cells [101]. However, in accordance with the sequential or combinatorial nature of cancer treatment, multiple therapeutic agents are often employed to eradicate malignant tumours. To this end, tubulin inhibitors are often administered following the administration of anthracyclines, often involved in cardiovascular adverse events, as described above. Therefore, the precise effect of tubulin inhibitors on adverse cardiac reactions yet remains open for debate [102].

2.2.1.2 Tyrosine kinase inhibitors

Tyrosine kinase inhibitors were once thought to be revolutionary for cancer treatment, providing targeted therapeutic replacements for other nonspecific and side-effect intensive interventions such as radiotherapy [103]. However, despite reduced rates of side toxicity, some tyrosine kinase inhibitors have been associated with cardiovascular system damage. Cardiotoxicity is linked primarily to their inhibition of major signalling pathways responsible for cardiomyocyte survival and maintenance, such as in the case of sorafenib or

vandetanib, which are VEGF signalling inhibitors [74,104–106]. Lapatinib, an inhibitor of EGFR signalling that targets both HER2 and EGFR, presents however a low level of cardiotoxicity. As previously discussed, the low level of cardiotoxicity of lapatinib is presumably due to its activation of the AMPK signalling pathway, which mobilises cardiomyocytes and increases ATP synthesis and storage [80]. In contrast, consider the tyrosine kinase inhibitor sunitinib, whose targets include VEGFR. Unlike lapatinib, sunitinib inhibits the AMPK signalling pathway and potentially inhibits energy metabolism and the PDGF signalling, all involved in the cardiomyocyte mechanical stress response [74,107]. Imatinib inhibits the chimeric oncogene (bcr-abl), the protein constructed by the fusion of a breakpoint cluster region (bcr) with an Abelson tyrosine kinase (abl). The cardiotoxicity of imatinib is primarily linked to the release of B-cell lymphoma 2 (Bcl-2) proteins, which produce mitochondrial damage [74]. Interestingly, it has been also shown that imatinib is able to inhibit the proliferation and apoptosis of neoplastic cells via inhibition of the hERG-encoded channel. *In vitro* applications of sunitinib and imatinib have demonstrated their capacity to activate CaMKII expression and activity, but with no effect on myocardial contractility [57]. In addition, sunitinib and imatinib demonstrate induction of high ROS levels, leading to an overall reduction of cell viability [108]. Other bcl-abl inhibitors, such as dasatinib and nilotinib, are also known to induce cardiac adverse events [109], which are presumably linked to the inhibition of VEGF signalling [110].

2.2.1.3 Monoclonal antibodies

In the context of antineoplastic immunotherapy, monoclonal antibodies represent one of the most promising oncological interventions. Monoclonal antibodies, identical to antibodies that are produced by the immune system, specifically bind to extracellular and cell-surface proteins, activating cellular apoptosis and blocking tumour proliferation. Therapy side effects are primarily linked to their target functions. HER2 inhibition by monoclonal antibodies is principally associated with cardiac dysfunction [80,111]. The toxicity of monoclonal antibodies is less predictable than for anthracyclines, although it

is generally reversible. As a representative example, trastuzumab is also known to downregulate Neuregulin-1, a signalling molecule involved in cardiac homeostasis and development [112]. The most common side effect of bevacizumab, an inhibitor of VEGFA, is hypertension, but myocardial infarction may occur in some cases [113]. Due to its inhibition of VEGFR signalling, bevacizumab also negatively affects the coagulation system.

2.2.1.4 HDAC inhibitors

Histone deacetylase (HDAC) inhibitors have long been used to treat neurological disease, but have recently been proposed as a powerful new class of antineoplastic agents. However, the growing concern regarding the cardiac safety of HDAC inhibitors has slowed down their progress in clinical trials. For example, romidepsin, which was approved for clinical use in 2009, has been suggested to produce a broad spectrum of cardiac adverse effects, including sudden cardiac death [114]. The cardiac adverse events potentially caused by romidepsin include QT prolongation and Torsade de Pointes arrhythmias, which are primarily linked to its effects on cellular trafficking of the hERG channel rather than a direct channel blockage [45,115]. To date, 17 HDACs exist, classically divided into 4 classes. However, cardiotoxicity is not a class specific characteristic. HDAC inhibitors of classes I, II, and IV are generally used as antineoplastics, while those suppressing class III (also known as sirtuins) could potentially be used to protect cardiac health by mainly reducing risk of thrombosis, atherosclerosis, and endothelial dysfunction [116,117].

2.2.1.5 Other antineoplastic agents

Bortezomib is a proteasome inhibitor approved for myeloma treatment that inhibits the ability of malignant cells to escape apoptosis. Potential side effects include neutropenia, thrombocytopenia and cardiotoxicity. Cardiac adverse events are considered reversible, however bortezomib has been reported to cause arrhythmias and to even lead to heart failure [118,119]. The

mechanism of cardiotoxicity is linked to its primary target, the ubiquitin-proteasome system, essential for cardiomyocyte and mitochondria function, and which also controls hERG channel trafficking [120].

In addition, recent reports have highlighted concerns over the cardiac safety of immune checkpoint inhibitors, newer and highly effective anticancer therapies [121,122]. Given their target toxicity, checkpoint inhibitors have been generally associated with immune-related complications, such in the case of vitiligo [123]. However, lethal myocarditis has also been reported during checkpoint inhibitor treatment, although the underlying mechanism of cardiotoxicity still remains unknown.

Used in leukemia treatment, arsenic trioxide is another example of a cancer drug that can produce undesired cardiac complications, such as ventricular and supraventricular tachycardias [124–126]. Inhibition of hERG trafficking induced by the drug results in IKr channel deficiency and action potential prolongation, in both animal and human ventricular cells [127].

2.2.2 Central nervous system agents

The electrophysiological function of both myocardial and nervous tissues require the propagation of action potentials, a process triggered by external and intracellular mechanisms, which involves the depolarisation and repolarisation of cellular membranes. Given the similarities in electrophysiology between both systems, central nervous system (CNS) agents that target neuronal electrophysiology can also affect cardiac tissue. Consequently, drugs that cause neurotoxicity may also result in cardiotoxicity [128]. However, neural tissue has a lower tolerance threshold compared to cardiac tissue, leading to the manifestation of neurotoxicity at lower doses. In addition, some CNS agents have alternative mechanisms of cardiotoxicity that do not involve alterations of the electrophysiology.

2.2.2.1 Local anesthetics

Both heart and nervous tissues depend on properly functioning sodium channels for the triggering of the action potential. Local anaesthetics frequently block sodium channels, therefore directly affecting both cardiac and nervous tissues. The most cardiotoxic and neurotoxic local anaesthetic is cocaine, first used in 1884 [129,130]. The next generation of local anaesthetics, such as bupivacaine and lidocaine, was intentionally developed to overcome these toxicity issues. However, clinical evaluations demonstrated that bupivacaine (markedly more cardiotoxic than other local anaesthetics) causes hemodynamic and electrophysiological disturbance, which may result in hypoxia, arrhythmia, and even cardiac arrest. Lidocaine and ropivacaine have also been shown to cause dose-dependent adverse cardiac reactions [131–133]. Together with sodium channel block, an additional hypothesised mechanism of toxicity involves the interaction between local anaesthetics and mitochondrial membranes by increasing the permeability of cardiolipin, therefore interrupting function and potentially inducing apoptosis [68,69].

2.2.2.2 Antidepressants

A preponderance of the available evidence emphasises antidepressants as another class of CNS agents with significant cardiovascular side effects. Indeed, several antidepressants were withdrawn from the market or restricted for use due to cardiac adverse reactions. Tricyclic antidepressants and neuroleptics directly interact with sodium, calcium and potassium channels, and in case of overdose can cause fatal arrhythmias and hypotension [134]. The next generation of antidepressants has been reported to have fewer cardiac adverse effects and are considered to be lower risk, but they too inhibit cardiac sodium, calcium and potassium channels [135,136]. For example, the selective serotonin inhibitor fluoxetine inhibits both the inward slow calcium current (ICaL) and IKr currents in addition to its primary targets. At the same time, fluoxetine has been also shown to disrupt hERG trafficking [48]. As a result, long-term administration of fluoxetine may result in mild

bradycardia. Venlafaxine, another selective serotonin inhibitor, is potentially more arrhythmogenic and has been shown capable of blocking the inward sodium current (INa) in both humans and animal model organisms [137–139].

2.2.2.3 Antipsychotic

Psychiatric disorders are known to be accompanied by an increased risk of cardiovascular disease [140–142]. In addition, many antipsychotics have been found to cause QT prolongation and cardiotoxicity [143]. For example, clozapine, one of the most effective psychiatric treatments for schizophrenia, has been linked to serious cardiovascular complications such as myocarditis, an inflammatory cardiac muscle disease [144,145]. As a result, clozapine is generally considered to be a last resort for patients in which other antipsychotics are ineffective. However, the mechanism underlying clozapine-induced myocarditis is unknown. Another example is sertindole, which was removed from the market in 1998 [146] due to the increased risk of QT-prolongation and cardiac sudden death, primary linked to its high-affinity block of IKr [147]. It was however re-launched on the European Market in 2005 as a medication that requires cardiac monitoring [148]. Other antipsychotics used to treat schizophrenia and bipolar disorder, such as thioridazine, ziprasidone, and haloperidol, can prolong QT intervals, lead to Torsade de Pointes, or even cause sudden cardiac death [149]. Ziprasidone has been shown to block IKr, thioridazine blocks both late sodium current (INaL) and IKr, and haloperidol blocks IKr, INa, and the inward L-type calcium current (ICaL), but shows no effect on INaL [135,150]. Interestingly, ziprasidone and clozapine have also been found to block alpha-adrenergic receptors [151], which are involved in the regulation of calcium-cycling in cardiomyocytes [152].

2.2.2.4 Neurodegenerative disease agents

At present, neurodegenerative diseases such as Alzheimer's and Parkinson's disease are incurable, attracting immense interest in the development of new prospective therapies. Among them is pergolide, a dopamine receptor

antagonist with potential application in the treatment of Parkinson's disease. However, administration of pergolide has been associated with the emergence of fibrosis in multiple tissues, including the heart [153,154]. It is believed that this side effect results from pergolide activating serotonin HTR2B receptors and subsequently inducing the transforming growth factor beta (TGF-beta) pathway [155,156]. As a result, pergolide use was restricted due to increased risk of fibrotic valvular heart disease.

2.2.2.5 Other central nervous system agents

Methysergide and ergotamine, serotonin 5-HT receptor inhibitors that are used to treat migraines, are known to share cardiac adverse effects with pergolide and to cause cardiac fibrosis [157]. In addition, a number of appetite suppressants including dexfenfluramine, fenfluramine and phentermine, in a combination known as fen-phen [158], chlorphentermine [159], and benfluorex [160] were lifted from the market due to their associated cardiac fibrosis. The primary mechanism by which fibrosis induction occurs is believed to involve the activation of the TGF-beta pathway through 5-HT2B receptors. Another discontinued appetite suppressant drug, sibutramine [161], was however linked to myocardial infarction. The underlying mechanism of this side effect has not been fully described, but *in vitro* experiments demonstrated that sibutramine blocks the IKr channel [162,163].

Opioid analgesics have also been shown to cause significant changes in cardiac electrophysiology [164]. This led to the withdrawal of levomethadyl acetate in 2001 from the European market [165] and propoxyphene in 2010 [166]. Although, FDA reported that levomethadyl acetate has no safety concerns [167], it was also lifted from the US market in 2003 [168].

2.2.3 Genitourinary system agents

Muscarinic receptor antagonists, an anticholinergic agent, are a class of drugs that has been approved for treatment of incontinence. However, muscarinic

receptor antagonists have a well-known potential to affect cardiac function. Although generally considered to be safe, they may increase heart rate, prolong QT, and increase cardiovascular risk when used in combination with other medications [169]. Moreover, terodiline, a muscarinic receptor antagonist, was withdrawn from the market due to serious cardiac side effects, such as Torsades de Pointes and even cardiac arrest among older patients [170]. The toxicity of terodiline is linked to the inhibition of the hERG encoded channel and disruption of calcium cycling, with its cardiotoxic effects depending on concentration and genetic background [16]. The withdrawal of terodiline has since then raised concerns about potential cardiac adverse reactions by the whole group of muscarinic receptor antagonists.

2.2.4 Gastrointestinal agents

Cisapride, a drug that increases motility in the upper gastrointestinal tract, was designed to block serotonin 5-HT₄, but it also inhibits the hERG encoded channel, in addition to 5-HT_{3A} and 5-HT_{2A} serotonin receptors [171]. Its marked hERG inhibition produced cardiac arrhythmias and Torsade de Pointes, leading to the withdrawal of cisapride from the market in most countries [172]. Tegaserod, a non-selective 5HT₄ inhibitor, was also withdrawn by the FDA in 2007 despite absence of reports showing it is able to inhibit the hERG channel [173]. New versions of selective 5-HT₄ blockers, such as clobopride and mosapride, are considered cardiac safe [174].

Loperamide, an opioid antagonist approved for the treatment of diarrhoea symptoms, primarily works by blocking the mu-opioid receptor. In addition to its primary target, loperamide has been shown to inhibit voltage-gated calcium channels and may produce cardiotoxic effects in humans and model organisms [170,175–177]. In 2016, the FDA issued an announcement indicating that high doses of loperamide, or its misuse thereof, could cause serious cardiac adverse events, including lethal heart attacks [178].

Another group of gastrointestinal agents labelled as cardiotoxic are proton pump inhibitors, such as omeprazole. While the direct connection between omeprazole and cardiotoxicity has never been fully established and the mechanism of action remains unknown [179], it is possible it arises due to the production of ADMA, which interferes with NO synthesis and accelerates endothelial cellular senescence [56,180,181].

2.2.5 Antihistamines

H1-Antihistamines are systematically associated with cardiac adverse events. They work to reduce the severity of allergic responses by blocking histamine H1-receptors on the cell surface. The biological role of histamines includes participation in the immune response to pathogens and allergens by increasing the permeability of capillaries. Within the cardiovascular system, the stimulation of H1 receptors results in the constriction of coronary blood vessels, inducing a positive chronotropic effect. However, two second-generation antihistamines, terfenadine and astemizole, were withdrawn from the market due to QT prolongation and Torsades de Pointes [182], linked to their direct block of IKr in addition to their main targets [183]. Diphenhydramine, the first antihistamine introduced to the market, also demonstrates a QT prolongation signature but only upon overdose.

2.2.6 Anti-inflammatory agents

Non-steroidal anti-inflammatory drugs (NSAIDs) are a broad class of agents that target one or both prostaglandin synthases (COX-1 and COX-2), and hence display anti-inflammatory properties. NSAIDs side toxicities, such as gastrointestinal toxicity, are primarily due to COX-1 inhibition. To circumvent these adverse side effects, new generations of NSAIDs selectively target the COX-2 enzyme. However, both selective COX-2 inhibitors and high doses of nonselective COX inhibitors display cardiotoxic side effects. To this end, in 2004 the FDA withdrew the COX-2 inhibitor rofecoxib – one of the most widely used drugs ever withdrawn from the market. While the significant

cardiotoxicity of rofecoxib seems to be linked to its unique metabolism [184], in 2007 the FDA issued a non-approval letter for the selective COX-2 inhibitor etoricoxib due to cardiotoxic concerns. In response to these concerns, the FDA strengthened the safety warning on NSAIDs labels in 2015 [185]. Long-term administration of both selective COX-2 inhibitors, such as diclofenac, and non-selective COX-1/COX-2 inhibitors, such as ibuprofen, naproxen, and indomethacin, has been shown to increase the risk of cardiac arrest [186]. The mechanism by which selective COX-2 inhibitors induce cardiotoxicity is primarily linked to blocking prostacyclin synthase with no effect on TXA₂ synthesis, leading to an increased risk of thrombosis [187]. In contrast to antimetabolites that induce TXA₂ formation, increasing the risk of ischemia, evidence suggests that COX-2 inhibitors decrease the risk of ischemia due to their ability to reduce inflammation [188,189].

2.2.7 Anti-infective agents

Anti-infective agents are compounds with “selective toxicity” against pathogens, such as bacteria, viruses or other microorganisms. Anti-infective agents vary by the mechanism of action and side effects.

2.2.7.1 Antibiotics

Antibiotics are among the most prescribed drugs in the world and work primarily by disrupting bioactive processes of pathogens, such as cell wall construction, protein and nucleic acid metabolism and repair. Since human cells share some functions with prokaryotic pathogens, some antibiotics (such as anthracyclines) are used as antineoplastic agents. The cardiotoxicity of cytostatic antibiotic agents has been discussed above. Macrolides, antibiotics whose primary mechanism of action is the inhibition of bacterial protein synthesis via inhibition of prokaryotic ribosomal subunits, and widely used to treat respiratory infections, have also been shown to block hERG channels [190]. Macrolides such as erythromycin, azithromycin and clarithromycin are considered arrhythmogenic [191]. It also has been shown that macrolides

cause mitochondrial toxicity by inhibiting protein synthesis in mitochondria [192]. Fluoroquinolones is another group of antibiotics that affects hERG channel and could cause Torsades de Pointes [193]. Due to this effect, fluoroquinolones such as grepafloxacin and sparfloxacin were discontinued in use.

2.2.7.2 Antivirals

Alongside with antibiotics, which are sometimes referred to as antibacterial agents, antiviral agents are a class of drugs used to target pathogens, but may cause adverse reactions in human tissues. Viruses vary widely by their mechanism of host tissue infection and transfection (transfer of nucleic acids with genetic information into host cells), and antiviral compounds are usually only effective against one family of viruses. The process of designing safe antiviral drugs is complicated by the fact that viruses use host cell structures to replicate. One of the major concerns associated with the administration of antiviral drugs is mitochondrial toxicity in liver, skeletal muscle, and heart tissue [66]. Azidothymidine, an antiretroviral HIV treatment, has been shown to induce mitochondrial dysfunction, leading to mitochondrial fragmentation and impaired fusion-fission cycle [67]. Azidothymidine also inhibits mitochondrial DNA polymerase alongside its main target, reverse transcriptase [66]. Another antiviral agent recently reported as cardiotoxic is sofosbuvir, an inhibitor of RNA polymerase nonstructural protein 5B, used to treat hepatitis C. Although the drug exhibited a cardiac-safe profile during clinical trials, several cases of severe bradycardia were reported post-marketing [194]. However, the mechanism by which sofosbuvir induces cardiac adverse effects is still unclear.

2.2.7.3 Other anti-infective

Pentamidine, used to treat leishmaniasis, trypanosomiasis and pneumonia, has been associated with QT-prolongation and ventricular arrhythmias in case of intravenous treatment [195]. In vitro experiments showed that while being a

poor direct blocker of IKr at therapeutic concentrations [196], pentamidine inhibits trafficking of the channel causing action potential prolongation in both animal and human cells [38,197].

2.2.8 Cardiovascular agents

Worsening of arrhythmic risk is one of most serious side effects of some antiarrhythmic agents. A number of highly potent drugs were lifted from the market or restricted in use due to their proarrhythmic effects, including dofetilide and encainide. Some non-selective calcium blockers, such as lidoflazine and prenylamine, were also suspended from the market due to their life-threatening QT prolongation. This side effect is primarily linked to the high-affinity block of the IKr channel by these drugs [198,199]. Buflomedil, an alpha-adrenoceptor antagonist, was withdrawn from the European market due to its unfavorable cardiac safety profile [200]. The selective beta-adrenoceptor blocker orciprenaline was removed from the market due to life-threatening cardiac side effects, linked to its primary target [201].

Landmark clinical trials were designed to investigate the efficacy and safety of cardiovascular agents, such as the Cardiac Arrhythmia Suppression Trial (CAST), which was focused on the class I antiarrhythmic agents encainide and flecainide [202]. Results of this study revealed that structural disease is the major risk factor for developing adverse reactions to flecainide. Later, the Survival With Oral d-Sotalol (SWORD) study raised concerns regarding the use of class III antiarrhythmic agents, and also demonstrated that patients after myocardial infarction are at a higher risk of developing drug-induced arrhythmias (Waldo, 1996). The same was shown for dronedarone, another class III antiarrhythmic compound, as reported in the Antiarrhythmic Trial with Dronedarone in Moderate to Severe CHF Evaluating Morbidity Decrease (ANDROMEDA) [203], and in the Permanent Atrial Fibrillation Outcome Study Using Dronedarone on Top of Standard Therapy (PALLAS) studies [204].

Although a cholesterol-lowering agent probucol has been associated with QT-prolongation and arrhythmias [195], it was discontinued from the US market in 2009 due to lack of efficacy against coronary artery disease [205]. Probucol affects hERG trafficking, and it has been consequently shown to cause electrophysiological abnormalities in neonatal rat ventricular cardiomyocytes [43].

2.2.9 Other agents

One of the negative effects of anti-diabetic drugs is their increased risk of heart failure, which has been demonstrated by a number of studies [206], including the famous Framingham Study [207]. In light of such a risk, in 2010 EMA and FDA issued prescribing and dispensing restrictions for different drugs containing rosiglitazone, a thiazolidinedione class anti-diabetic drug [208]. However, after a thorough examination of large clinical trials, the FDA removed those restrictions and eliminated the risk evaluation and mitigation strategy for rosiglitazone, confirming its cardiac safety [209]. In addition to the debatable safety of rosiglitazone, concerns regarding the safety of saxagliptin and alogliptin were raised [210].

In 2007, the cough suppressant clobutinol was lifted from the market by its manufacturer due to risk of QT prolongation [211]. Animal experiments suggested that clobutinol inhibits the IKr channel and induces Torsades de Pointes [212].

2.3 Conclusions

In this chapter, I have presented a comprehensive discussion on drug-protein mechanisms underlying drug-induced cardiotoxicity in humans. This significantly broadens other studies by covering side effects and cardiotoxic mechanisms of cardiovascular, anticancer, central nervous system, genitourinary, gastrointestinal, antihistamines, anti-inflammatory, and anti-infective agents, beyond their direct interactions with cardiac

electrophysiology. As a result, I demonstrate that many of these drug classes share common modes of action on cardiac function and on adverse cardiac events, with cardiotoxicity frequently resulting from the simultaneous interruption of key myocardial functions and viabilities (mitochondrial dysfunction, inhibition of major signalling pathways responsible for cardiomyocyte survival and maintenance, fibrosis, NO synthesis, calcium cycling and cellular trafficking).

The understanding of the mechanisms of drug-induced toxicity and differences in cardiac safety profiles of therapies is a subject of great importance for the development of new compounds, and of safety assays for preclinical testing. In the same fashion, a refined understanding of the multifaceted components of cardiotoxicity can lead to the development of cardioprotective agents, with the potential of reducing the manifestation and damage of otherwise effective but risky drugs. As an exemplar, statins have been proposed to lower anthracycline-induced harm via a reduction of ROS initiated signalling and regulation of TOP2B [213]. Equivalently, the activation of the AMPK pathway by lapatinib may underlie its lower cardiotoxicity compared to other tyrosine kinase inhibitors. In fact, other AMPK activators, such as metformin, have been shown to be protective for several heart conditions [214].

As stated above, this chapter focuses on mechanisms of drug-induced cardiac side reactions, concentrating its scope on drug-protein interactions. Drug-drug and drug-food interactions are also known to promote adverse cardiovascular effects, usually linked to competition between enzymes participating in drug metabolism, or alterations in drug metabolism [215,216]. In addition to these, recent advances in the annotation of RNA-RNA interactions suggest they might be of even higher importance in cardiac disease, such as heart failure and hypertrophy [217,218]. It has also been proposed that RNA-RNA interactions are involved in the regulation of channel trafficking, and dysregulation of these interactions could cause arrhythmia [219]. Among those discussed in this review, the anti-cancer drugs 5-FU and

cisplatin are also known to interact with micro-RNA and perturb RNA-RNA intercommunication [220].

All these mechanisms of drug-induced cardiotoxicity therefore constitute important and promising prospects for future investigations into drug discovery, safety pharmacology, and continuous pharmacovigilance studies. As a result, their integration as a priority into multi- and inter-disciplinary approaches across academia, industry, and healthcare settings, for the complex characterisation of human heart physiology [221], is therefore expected to yield major advances for the analysis and prediction of adverse drug reactions and cardiotoxicity in human, both in health and disease.

Table 2-1 List of compounds with their targets and cardiac side effects.

Drug	Introduced Withdrawn (Years)	- Mechanism of action with targets	Side effects on cardiac function	Mechanism of cardiac toxicity	Ref.
Antineoplastic agents					
Doxorubicin	1978 - NA	DNA intercalation, inhibition of TOP2B	Arrhythmias, heart failure	Mitochondrial toxicity, oxidative stress leading to apoptosis	[87,222]
Daunorubicin	1980 - NA	Inhibition of TOP2B	Arrhythmias, heart failure	Mitochondrial toxicity, oxidative stress learning to apoptosis	[222,223]
Idarubicin	1989 - NA	Inhibition of TOB2B	Arrhythmias, heart failure	Mitochondrial toxicity, oxidative stress learning to apoptosis	[222]
5-fluorouracil	2000 - NA	DNA cross-linking	Arrhythmias, myocardial ischemia, heart failure	TXA ₂ activation leading to cardiac remodelling, electrolyte imbalances	[91,224]

Cisplatin	1978 - NA	DNA cross-linking	Arrhythmias, myocardial ischemia, heart failure	TXA ₂ activation leading to cardiac remodelling, electrolyte imbalances	[91,225]
Paclitaxel	1995 - NA	Inhibition of beta subunit of tubulin, Bcl-2	Bradycardia, myocardial ischemia, heart failure	Apoptosis of cardiac endothelial cells	[226,227]
Docetaxel	1991 - NA	Inhibition of beta subunit of tubulin, Bcl-2	Bradycardia, myocardial ischemia, heart failure	Apoptosis of cardiac endothelial cells	[100,228]
Vinblastine	1960s - NA	Inhibition of alpha beta and delta subunits of tubulin	Myocardial ischemia, heart failure	Apoptosis of cardiac endothelial cells	[99]
Cytarabine	1966 - NA	DNA intercalation, Inhibition of DNA	Bradycardia, heart failure,	Unknown	[97,98]

		polymerase beta	ischemia		
Imatinib	2016 - NA	Bcr-Abl kinase, KIT, RET	Systolic heart failure, left ventricular dysfunction	heart heart left	Release of Bcl-2 proteins leading to mitochondrial toxicity, oxidative stress leading to apoptosis, disruption of cardiomyocytes survival via VEGF signalling inhibition [109]
Dasatinib	2016 - NA	Bcr-Abl kinase, EPHA2, LCK, YES1	Heart failure		Disruption of cardiomyocytes survival via VEGF signalling inhibition [109]
Sorafenib	2006 - NA	Inhibition of BRAF kinase, FLT3, RAF1, VEGFR3, VEGFR2, PDGFRB, KIT, FGFR1, RET, FLT1	Heart failure, myocardial ischemia, QT prolongation	Heart failure, myocardial ischemia, QT	Disruption of cardiomyocytes survival via VEGF signalling inhibition [229]
Lapatinib	2004 - NA	Blocking of EGFR,	Left ventricular		Disruption of cardiomyocytes [230]

		HER2	ejection fraction, survival via EGF signalling congestive heart inhibition failure	
Nilotinib	2005 - NA	Inhibition of ABL1, blocking of KIT	Myocardial ischemia	Disruption of cardiomyocytes [109] survival via VEGF signalling inhibition
Sunitinib	2006 - NA	Inhibition of PDGFRB, KIT, FLT3, PDGFRA	Long QT, left ejection fraction, myocardial infarction	Disruption of cardiomyocytes [107] survival via VEGF signalling inhibition, disruption of stress response via PDGF signalling inhibition, disruption of energy homeostasis and mitochondrial fusion-fission system via inhibition of AMPK signalling
Vandetanib	2011 - NA	Inhibition of VEGFA, EGFR, PTK6, TEK	Long QT	Disruption of cardiomyocytes [230,231] survival via VEGF signalling inhibition, disruption of

				cardiomyocytes survival via EGF-signalling inhibition	
Trastuzumab	2016 - NA	Blocker of HER2	Heart failure, tachycardia	Disruption of cardiomyocytes survival via EGF signalling inhibition	[80,112]
Bevacizumab	2016 - NA	Blocker of VEGFA	Heart failure	Disruption of cardiomyocytes survival via VEGF signalling inhibition	[113]
Bortezomib	2003 - NA	Inhibition of PSMB5, PSMB1	Heart failure, arrhythmia	channel trafficking inhibition	[118,119]
Romidepsin	2009 - NA	Inhibition of HDAC1, HDAC2, HDAC4, HDAC6, ABCC1	QT-prolongation, myocardial infarction	channel trafficking inhibition	[232]
Ipilimumab	2011 - NA	Inhibition of CTLA-4	Lethal myocarditis	Unknown	[121,122]
Nivolumab	2017 - NA	Inhibition of PD-1	Lethal	Unknown	[121,122]

myocarditis

Arsenic trioxide	2000 - NA	Inhibition of TXNRD1, activation of IKBKB, JUN, MAPK3, MAPK1	QT-prolongation, channel trafficking inhibition, tachicardia	[124–126]
------------------	-----------	--	--	-----------

Anti-inflammatory agents

Rofecoxib	1999 - 2004	Inhibition of COX2	Myocardial infarction	Blocking prostacyclin synthase	[233]
Diclofenac	1986 - NA	Inhibition of COX1, COX2, SCN4A, ASIC1 and potentiation of ALOX5	Myocardial infarction	Blocking prostacyclin synthase	[186,234]
Naproxen	1978 - NA	Inhibition of COX1, COX2	Myocardial infarction	Blocking prostacyclin synthase	[186]

Ibuprofen	1978 - NA	Inhibition of COX1, COX2	Myocardial infarction, hypertension	Blocking prostacyclin synthase	[186,234]
Indomethacin	1966 - NA	Inhibition of COX1, COX2, PLA2G2A, GLO1 and activation of PPARG	Myocardial infarction	Blocking prostacyclin synthase	[186,234]
Etoricoxib	2002 - 2007	Inhibition of COX2	Thrombotic events	Blocking prostacyclin synthase	[186,234]

Central nervous system agents

Lidocaine	1944 - NA	Inhibition of SCN10A, SCN9A, SCN5A, blocking of EGFR	Bradycardia, cardiac arrest,	Inhibition of voltage gated sodium channel, mitochondrial toxicity	[133,235]
Cocaine	1884 - NA	Inhibition of SLC6A3, SLC6A2,	Left ventricular hypertrophy,	Inhibition of voltage gated sodium channel, mitochondrial toxicity	[130]

		SLC6A4, SCN5A	arrhythmias		
Bupivacaine	1965 - NA	Inhibition of SCN10A	Ventricular of arrhythmias, myocardial depression	Inhibition of voltage gated sodium channel, mitochondrial toxicity	[132]
Fluoxetine	1990 - NA	Inhibition of SLC6A4	Bradycardia	Inhibition of ICa-L and IKr, channel trafficking inhibition	[236]
Venlafaxine	1986 - NA	Inhibition of SLC6A4, SLC6A2	QT prolongation, arrhythmias	Inhibition of INa	[237,238]
Ziprasidone	2001 - NA	Inhibition of DRD2, HTR2A, DRD1, DRD5, DRD3, DRD4, HTR1B, HTR1D, HTR1E, HTR2C, HTR3A, HTR6, HTR7, HRH1, ADRA1A,	QT prolongation, TdP, Sudden cardiac death	Inhibition of IKr	[149]

		ADRA1B, ADRA2A, ADRA2B, ADRA2C, CHRM1, CHRM2, CHRM3, CHRM4, CHRM5 and activation of HTR1A,		
Haloperidol	1967 - NA	Blocking of DRD2, QT prolongation, GRIN2B, DRD1 and TdP, sudden inverse activation of cardiac death DRD3	Inhibition of IKr, INa, ICa-L	[149]
Thioridazine	1978 - NA	Blocking of DRD2, QT prolongation, DRD1, ADRA1A, TdP, sudden ADRA1B, HTR2A, cardiac death inhibition of KCNH2	Inhibition of INa-L, IKr	[149]
Clozapine	1991 - NA	Blocking of DRD2, Myocarditis, HTR2A, DRD1, cardiomyopathy	Unknown	[144,145]

		DRD3, DRD4, HTR1A, HTR1B, HTR1D, HTR1E, HTR2C, HTR3A, HTR6, HTR7, HRH1, HRH4, ADRA1A, ADRA1B, ADRA2A, ADRA2B, ADRA2C, CHRM1, CHRM2, CHRM3, CHRM4, CHRM5			
Sertindole	1998 ⁽¹⁾ - 2014	Blocking of DRD2, QT prolongation, HTR2A, HTR2C, TdP, sudden HTR6 cardiac death	Inhibition of IKr	[143]	
Pergolide	1987 - NA	Blocking of DRD2, DRD3, DRD4, Valvular heart HTR1A, HTR2B, disease HTR1D, ADRA2B,	Induction of fibrosis via HTR2B- induced activation of TGF-beta signalling	[154,155]	

		ADRA2A, ADRA2C, DRD5, DRD1, HTR2C, ADRA1B			
Methysergide	1965 - NA	Blocking of HTR2B, HTR2C, HTR2A, HTR7, activation of HTR1A and binding HTR1B, HTR1F, HTR1E	Valvular disease	heart	Induction of fibrosis via HTR2B- induced activation of TGF-beta signalling [154]
Ergotamine	1925 - NA	Activation of HTR1D, HTR1B, HTR2A, DRD2, ADRA1A	Valvular disease	heart	Induction of fibrosis via HTR2B- induced activation of TGF-beta signalling [157]
Dexfenfluramine	1996 - 1997	Inhibition of SLC6A4	Valvular disease	heart	HTR2B-induced activation of TGF- beta signalling [239,240]
Fenfluramine	1973 - 1997	Inhibition of SLC6A4 and blocking of	Valvular disease	heart	HTR2B-induced activation of TGF- beta signalling [240,241]

HTR2B

Phentermine	1959 - 1997	Inhibition of SLC6A2, SLC6A4, SLC6A3, blocking of MAOB, MAOA	Valvular disease	heart	HTR2B-induced activation of TGF- beta signalling	[240]
Chlorphentermine	1966 - 1969	Blocking of 5-HTs	Pulmonary disease	heart	HTR2B-induced activation of TGF- beta signalling	[242]
Benfluorex	1972 - 2009	Blocking of 5-HT2B	Valvular disease	heart	HTR2B-induced activation of TGF- beta signalling	[243]
Sibutramine	2001 - 2002	Inhibition of SLC6A4, SLC6A2, SLC6A3	Myocardial infarction		Inhibition of IKr	[244–246]
Propoxyphene	1957 - 2010	Activation of OP1, OP2, and OP3	QT prolongation, TdP		Inhibition of IKr	[164]
Levomethadyl	1991 - 2003	Activation of	QT prolongation,		Inhibition of IKr	[247]

acetate		OPRM1	TdP	
Gastrointestinal agents				
Loperamide	1976 - NA	Blocking of OPRM1, OPRK, POMC modulation of CALM1	Cardiac arrest, QT prolongation, ventricular tachycardia, TdP	Inhibition of voltage gated calcium channels [248,249]
Omeprazole	1987 - NA	Inhibition of the ATP4A	Acute myocardial infarction, heart failure	Disruption of NO synthesis via ADMA production [179]
Cisapride	1980 - 2000	Blocking of HTR4, HTR3A, HTR2A and inhibition of KCNH2	Ventricular arrhythmia, QT prolongation, TdP, cardiac arrest	Inhibition of IKr [250]

Tegaserod	1997 – 2007 ⁽²⁾	Blocking of HTR4, HTR2B, HTR2C, HTR2A	Ischemia	Inhibition of IKr	[173]
Genitourinary system agent					
Terodiline	1975 - 1991	Blocking mAChRs	Ventricular tachycardia, cardiac death	Inhibition of IKr, block of calcium- cycling	[170,251]
Anti-Allergic agents					
Astemizole	1992 - 1999	Blocking of HRH1 and inhibition of KCNH2	long QT syndrome, TdP	Inhibition of IKr	[182,252]
Diphenhydramine	1946 - NA	Blocking of HRH1 and CHRM2	QT prolongation	Inhibition of IKr	[138,253]
Terfenadine	1985 - 1997	Blocking of HRH1	QT prolongation, TdP	Inhibition of IKr	[254]

 Anti-infective

			QT prolongation,		
Clarithromycin	1993 - NA	Inhibition of rplJ, SLCO1B1, SLCO1B3	myocardial infarction, arrhythmias, cardiac death	Mitochondrial toxicity	[255,256]
Azithromycin	1988 - NA	Inhibition of 23s rRNA, rplD, rplV, PADI4	QT prolongation, TdP, cardiac death	Mitochondrial toxicity	[257]
Erythromycin	1955 - NA	Inhibition of 23s rRNA, MLNR, KCNH2, ALB	QT prolongation, ventricular tachycardia, TdP, ventricular fibrillation	Mitochondrial toxicity	[191]
Grepafloxacin	1999 - 1998	Inhibition of gyrA, parC	QT prolongation	Inhibition of IKr	[193]

Sparfloxacin	1997 - 2001	Inhibition of parC, gyrA, TOP2A	QT prolongation	Inhibition of IKr	[193]
Azidothymidine	1989 - NA	Inhibition of Pol, TERT	Dilated cardiomyopathy	Mitochondrial toxicity	[258]
Sofosbuvir	2013 - NA	Inhibition of NS5b	Bradycardia	Unknown	[194]
Pentamidine	1975 - NA	Inhibition of RNA transfer	QT prolongation, arrhythmias	Inhibition of IKr trafficking	[195]

Cardiovascular agents

Dofetilide	2000 - 2004	Inhibition of KCNH2, KCNK2, KCNJ12	QT prolongation, TdP	Inhibition of IKr	[259]
Encainide HCL	1987 - 1991	Inhibition of SCN5A	QT prolongation, TdP	Inhibition of IKr	[202]
Prenylamine	1960s - 1988	Blocking of MYLK2,	QT prolongation,	Inhibition of IKr	[260]

		CALM	TdP, cardiac sudden death, ventricular tachycardia		
Lidoflazine	1973 - 1989	Blocking of calcium channels	QT prolongation	Inhibition of IKr	[261]
Buflomedil	1970s ⁽³⁾ - 2011	Blocking of ADRA1A, ADRA2A	QT prolongation, cardiac arrest	Unknown	[260,262]
Orciprenaline	1965 - 2011	Activation of ADRB2	Tachycardia, palpitations	Activation of ADRB2	[201]
Probucol	1995 – 2009 ⁽⁴⁾	Inductor of LDL catabolism	QT prolongation, arrhythmias	Inhibition of IKr trafficking	[195]
Other agents					
Rosiglitazone	1994 - NA	Activation of PPARG	Heart failure ⁽⁵⁾	Unknown	[263]
Saxagliptin	2009 - NA	Inhibition of	Heart failure	Unknown	[264]

		DPP4			
Alogliptin	2013 - NA	Inhibition of DPP4	Heart failure	Unknown	[265]
Clobutinol	1963 - 2007	Inhibition of GABA receptors	QT prolongation	Inhibition of IKr	[211]

Drug targets with a known pharmacological action of drugs were extracted from DrugBank database (www.drugbank.ca). Information on the year a drug was launched on market and withdrawn from the market was collected from MeSH database (<https://www.ncbi.nlm.nih.gov/mesh>) and from literature.

- (1) Australia approved
- (2) Available only under a restricted distribution program
- (3) EMA approved only
- (4) Removed from the US market
- (5) FDA lifted restrictions for rosiglitazone and confirmed its safety

3 Computational approaches for prediction and investigation of drug-induced cardiac side effects

3.1 Introduction

In this chapter, I begin with a review of existing studies on transcriptional markers of drug-induced cardiac damage. Following this, I provide an introduction to gene expression analysis. I then introduce and review the methods of predicting adverse reactions caused by drugs using machine learning methods. I provide a summary of recent case studies describing machine-learning-based approaches to predicting drug-induced cardiotoxicity. This is followed by a review of the known limitations and challenges of machine learning analysis. After, I consider current literature on cardiac electrophysiological modelling and simulation, introducing populations of models and the set of biomarkers I will use to evaluate those model outputs. Finally, I provide an overview of the available databases for the modelling of drug-induced cardiotoxicity.

3.2 Gene expression signatures of drug-induced cardiac damages

The accumulated body of evidence suggests that gene expression signatures could be used as biomarkers of cell response to drugs [266–268]. One of the studies demonstrated that 35 out of 84 genes involved in cardiac function and stress response were associated with single and repeated incubations of iPCS-derived cardiomyocytes with three anthracyclines: doxorubicin, daunorubicin and mitoxantrone [269]. Anthracyclines are known for their dose-related and cumulative cardiotoxicity, and this finding suggests that severe cardiotoxicity could be predicted at the preclinical level.

Multiple studies have explored transcriptome perturbations in cells exposed to doxorubicin at different doses and incubation times [270–272]. In 2015, Holmgren and colleagues [272] reported results for 48 hours of incubation and

post-exposure (12 days after doxorubicin washout) for cardiomyocytes derived from human embryonic stem cells (hESC), demonstrating the activation of cellular response to stress, including apoptosis and p53 signalling pathways, master regulators of apoptosis and cell cycle arrest proteins. Notably, levels of lactate dehydrogenase and cardiac-specific troponin T, markers that are usually measured in the clinic to evaluate cardiac damage, were only predictive alone of an acute response. In addition, gene expression changes were more specific with respect to the p53 signalling network, and the growth differentiation factor 15 (GDF15) showed constant upregulation and a dose-specific response after exposure and in post-treatment. Gene expression and pathways analyses performed using data from iPSC-derived cardiomyocytes, incubated chronically (48 hours) with doxorubicin and in post-exposure cells (5 days incubation after doxorubicin washout), showed similar results [271]. After 48 hours of incubation, cells showed downregulation of cell cycle progression and cell division pathways and upregulation of apoptotic pathways. Interestingly, expression values of some genes of post-exposure cells demonstrate a return to the baseline values.

iPSC-derived cardiomyocytes were also used to explore the transcriptional response to HDAC inhibitors [273]. To analyse the time evolution of the response, the authors compared 2 and 4 hours and 12 and 24 hours of incubation of a severe cardiotoxic group (0.5 μ M dacinostat, 0.1 μ M panobinostat, and 5 μ M vorinostat), an intermediate group (0.01 μ M panobinostat and 0.3 μ M vorinostat) and a “relative” noncardiotoxic group (0.3 μ M entinostat, 1 μ M tubastatin-a, and 0.1 μ M vorinostat). Both the severe and intermediate groups demonstrated similar time-dependent dynamics suggesting a common mechanism of action. Adrenergic receptor activity microtubule and cytoskeleton-based transport, cardiac muscle cell contraction, and Z-disc and unfolded protein binding were among the top most-perturbed expression pathways, which is in agreement with the observed contractile dysfunction of cells caused by drugs.

Tyrosine kinase inhibitors are yet another pharmacological group used to determine the value of gene expression signatures in predicting drug-induced cardiac disorders. In this way, Wang et al. [274] profiled various doses and incubation times of sunitinib, sorafenib, lapatinib and erlotinib in iPSC-derived cardiomyocytes, and showed that their transcriptional signatures are diverse and different from those of doxorubicin. Nonetheless, the authors found that the response of iPSC-derived cardiomyocytes to tyrosine kinase inhibitors is more complex and that cells showed little or no cell death, which had been previously suggested as a potential mechanism of toxicity. They also reported that gene sets related to metabolism, lipid biosynthesis and extracellular remodelling were perturbed by all four tyrosine kinases.

Earlier, Matsa and colleagues proposed to use iPSC-derived cardiomyocytes to evaluate potential patient-specific responses to drugs [275]. They first analysed individual signatures of untreated cells derived from seven patients and then explored how those cells responded to the treatment of tacrolimus and rosiglitazone, which are known to cause cardiac complications in patients.

3.3 Gene expression analysis

Gene expression analysis is routinely used in biomedical research. Gene expression profiling has evolved from the detection of expression levels of single genes to genome-wide analysis using microarray systems or RNA sequencing. Moreover, the challenge no longer lies in obtaining mRNA levels but in interpreting the results. Typically, gene expression levels are profiled for samples of interest and matching controls, which are then presented in tabular format, i.e. as a matrix $i \times j$, where i is the number of genes and j is the number of samples. The number of profiled samples is usually small, while the number of profiled genes or transcripts is large, so gene expression suffers from the curse of dimensionality. As a result, several approaches for gene expression analysis exist: differential gene analysis and gene sets analysis. Differential gene analysis operates with individual genes in order to obtain the list of the most perturbed genes, and then link them to their

potential biological functions. On the other hand, gene sets (also called pathway analysis) works with lists of genes united by their function, location in the cell, etc.

3.3.1 Gene list annotation

Historically, the difference between the control sample and the examined samples is calculated as fold change (FC), so the differential expression of a gene is a relative value:

$$(3.1) FC_A = \frac{\mu(A_t)}{\mu(A_c)},$$

where $\mu(A_c)$ is mean of the expression level of gene A in the control samples and $\mu(A_t)$ in the sample of interest (treated or diseased), respectively. For ease of manipulation, fold changes are usually $\log_2$ transformed. For example, if the $\log_2FC$ for a particular gene equals 3, it means that the sample of interest demonstrates an eight-fold increased expression of this gene. In contrast, if the $\log_2FC$ equals -3, it means that the expression of a gene is eight-fold decreased. The threshold of significance of $\log_2FC$ is somehow arbitrary, and usually a $|\log_2FC| \geq 1.5$ is considered to be meaningful. However, this choice can increase false results, especially when the number of samples is relatively small. As a result, several more robust metrics have been proposed to measure changes in gene expression, such as Pearson correlation, logistic prediction, or Gaussian separation scores [276,277]. Yet, these metrics are complicated in implementation, so FC is still dominant in gene expression analysis.

The main statistical approach to detect differential expression is the t-test. However, t-tests have high false-positive rates, especially on small sample sizes [278]. To address these problems, several pipelines have been introduced, which incorporate different normalisation steps in order to reduce bias and increase statistical power [279–281]. As an example, DESeq, a popular package for RNAseq analysis, performs normalisation on the expression of the strength parameter, calculated for each sample as:

$$(3.2) \hat{s}_j = med_i \frac{A_{ij}}{\prod_{z=1}^m A_{iz}^{1/m}},$$

$$(3.3) q_{ip} = \frac{1}{m_p} \sum_{j:p(j)=(p)} \frac{A_{ij}}{\hat{s}_j},$$

where A_{ij} is the expression level of gene i for sample j , A_{iz} is the expression level of gene i for sample z , m is the number of samples, med is for median, and $\hat{s}_j$ is the estimated size factor for sample j .

3.3.2 Gene list functional annotation

One of the possible methods of biological interpretation of gene lists is functional annotation. The Gene Ontology (GO) consortium, launched in 1998, is one of the largest bioinformatics initiatives [282]. GO data have been automatically collected based on over 100,000 peer-reviewed scientific papers. GO contains gene functional descriptions and terms that are organised into three categories: cellular components, molecular functions, and biological processes. For example, the terms for the human TOPMT1 gene, which encodes the mitochondrial topoisomerase I, are the following: cellular components are nucleus, chromosome, nucleolus, mitochondrion, replication fork protection complex, mitochondrial nucleoid; molecular functions are DNA binding, DNA topoisomerase type I activity; biological processes are DNA replication, DNA topological change, chromatin remodelling and chromosome segregation.

However, single-gene analysis always complicates the biological interpretation of the results and reproducibility. For example, two studies of the same biological system may show no overlap between their list of statistically significant differentially expressed genes [283]. In order to increase the robustness of gene expression signatures and to overcome the statistical challenges of gene expression analysis, a number of integrative approaches that manipulate gene sets and could be presented as pathways or gene networks have been proposed. These approaches could be roughly divided into three groups: gene enrichment analysis, gene network analysis, and methods that combine characteristics from both a gene enrichment analysis and a gene network analysis.

Gene enrichment analysis treats genes as part of a gene list that constructs a pathway and calculates the enrichment score for a particular pathway. One of the most popular gene enrichment approaches is the Gene Set Enrichment Analysis (GSEA) [34], which calculates an enrichment score to show the representation of highly perturbed genes in a particular gene set. The statistical significance of the enrichment score is calculated using a permutation test procedure:

$$(3.4) z_k = \frac{1}{\sqrt{A}} \sum_{k=1}^m t_k \times I_{km},$$

where I_{km} is an incidence matrix showing presence of a gene m in a pathway k , while t_k is a per-gene test statistics for a pathway k .

GSEA also allows computing robust z-scores for gene sets, which are normalized on the statistics from (3.4). Z-scores are usually calculated for pathways with more than 10 genes in them. In contrast, the limitations of GSEA include the use of Kolmogorov-Smirnov statistics, which are known for lack of sensitivity [35]. In addition, this entire class of methods that analyses gene sets neglects information on pathway topology and the hierarchy of signal transduction, which means that changes in the expression of core regulators from the top of the pathway hierarchy should biologically impact the whole pathway more than changes in the expression level of elements at the bottom of the pathway.

As an alternative, gene network analysis represents a gene list as a directed or undirected graph of interactions. One of the most popular approaches for gene network analysis is the signalling pathway impact analysis (SPIA) [36]. SPIA calculates two scores: the perturbation factor (PF) for each gene, and a total net perturbation (t_A) for each pathway:

$$(3.5) PF = \log_2 FC(A_i) + \sum_{j=1}^m \beta_{ij} \cdot \frac{PF(A_j)}{Nds(A_j)},$$

$$(3.6) Acc_i = PF(A_i) - \log_2 FC(A_i),$$

$$(3.7) t_A = \sum_{i=1}^m Acc(A_i),$$

where $\log_2 FC(A_i)$ is a log-transformed fold change value of a gene i , β_{ij} is a weight coefficient, $PF(A_j)$ is a perturbation factor for an upstream gene, $Nds(A_j)$ is the number of downstream genes, m is the number of genes in the pathway, and $Acc(A_i)$ represents the accumulation value for gene i . Weight coefficients quantify the strength of the interaction between elements and are usually set so that $\sum_{j=1}^m \beta_{ij} = 1$.

The third group of pathway analysis tools are based on pathway topology, and are used to assign activation or repression roles to nodes and genes in the pathway. The overall activation or suppression of the pathway is then calculated. Recently published, the in-silico Pathway Decomposition Network Analysis (iPANDA) [284] is an algorithm that calculates the pathway activation scores (PASSs) for each pathway based on $\log_2 FC$ as a metric of differential changes in expression. iPANDA also operates with nodes as a means of combining genes. These scores are given by:

$$(3.8) G_{ip} = w_i^s \cdot w_{ip}^T \cdot B_{ip} \log_2 FC(A_i),$$

$$(3.9) M_{ip} = \max(w_i^s) \cdot \frac{1}{N} \sum_i^N (w_{ip}^T \cdot B_{ip} \log_2 FC(A_i)),$$

$$(3.10) iPANDA_p = \sum_i G_{ip} + \sum_j M_{jp},$$

where $\log_2 FC(A_i)$ is a log-transformed fold change value of a gene i , B_{ip} is the activation sign (which equals to 1 or -1 based on the gene i contribution of the pathway p), w_i^s is the statistical weight coefficient, w_{ip}^T is the topological weight coefficient for gene i in pathway p , and N is a number of genes in a module. G_{ip} represents the contribution of individual genes and M_{jp} is the contribution of gene modules.

However, gene enrichment and gene network analyses are supervised and heavily dependent on the pathway annotations. In addition, gene-network-based approaches also require information on the pathway topology. Pathway annotations could be obtained from pathway databases.

3.4 Applications of machine learning in cardiotoxicity prediction

Above, I have discussed the conventional methods for transcriptomic data analysis used to identify the transcriptional changes in response to incubation with cardiotoxic drugs. At the same time, machine learning is gaining popularity in biological data analysis [285–288]. By providing methods for data analysis without a prior assumption of data dependencies, algorithms can learn without “being explicitly programmed”. Thus, a variety of machine learning algorithms could be applied to the supervised task of cardiotoxicity prediction.

3.4.1 Case studies

One interesting approach to assessing the cardiotoxicity prediction was proposed by Huang and colleagues in 2011 [289], who analysed the protein-protein interaction networks for drugs that were associated with adverse cardiac reactions in the SIDER database. The model was built using information on general cardiotoxicity for 578 drugs. However, while being reasonably accurate (their best model was support vector machines (SVM), with an area under the curve (AUC) of 0.771), their method was limited by the list of known drug targets and protein-protein interactions (so it could only be applied to well-described and characterised drugs), and it was only evaluated with ten-fold cross-validation.

A different machine learning study in combination with biophysically-detailed modelling and simulation of ventricular cells achieved a binary stratification of pro-arrhythmic drugs with a performance of 96.3% AUC on leave-one-out cross-validation [290]. The authors applied principal component analysis (PCA) to the action potential and the calcium transient measurements produced by simulations and then used the first four components, explaining most of the data variability as input features for SVM algorithms. However, the algorithm does not account for how the drugs could affect channel trafficking, which plays a role in the triggering of arrhythmias by some drugs.

Using transcriptional profiles and fingerprints of 251 drugs, Wang et al focused on the prediction of gastrointestinal, liver and kidney toxicities, and

myocardial infarction, a single form of cardiotoxicity, using an extra trees algorithm for multi-label classification [291]. The proposed algorithm used 166-bit Molecular ACCess System chemical fingerprints (MACCS) and Gene Ontology enrichment vectors as an input vector. It returned 1053 adverse reactions, including myocardial infarction. Interestingly, the algorithm demonstrated variable predictive accuracy across adverse reactions. The authors trained the model with three-fold cross-validation and achieved an AUC of 80%, yet no further testing was reported.

Messinis et al. [292] developed a transcriptomic-based predictor of drug-induced cardiomyopathy with 31 drugs. They first transformed gene expression drug responses into a set of signalling pathways using integer linear programming. Afterwards, they applied elastic net logistic regression to binary stratify drugs as toxic or not. While the reported accuracy of the approach is an AUC of 100%, the model was only tested with leave-one-out cross-validation.

The above described approaches for drug safety on the cardiovascular system represent the current state-of-the-art methods. It should be noted that most of these are limited by the type of cardiotoxicity they are able to detect, and the classes of drugs they are able to work with. The most comprehensive database of 578 drugs with various side effects extracted from the SIDER database was used to predict cardiotoxicity based on protein-protein interaction networks (Table 3-1). However, the SIDER database provides information on compounds that were reported by healthcare professionals, so the approach inherits all the limitations of the database. Importantly, although all these studies reported relatively good accuracies under different cross-validation strategies (random split of samples or leave-one-drug out), none of them conducted an independent testing on drugs previously unseen by the trained model. This is crucial to ensure the translatability of proposed approaches to real-world applications, thus representing an important limitation of previous work. In addition, none of the previous algorithms were developed to predict all major forms of drug cardiotoxicity.

Table 3-1 Summary of machine learning approaches to predicting cardiac disorders. AUC: area under the curve; SVM: support vector machines; PCA: principal component analysis; N: number; Ref: reference;

Cardiac disorder type	Best-performing algorithm	Performance (AUC, %)	Input data	N of investigated drugs	Ref.
General cardiotoxicity	SVM	77.1	Protein-protein interactions	578	[289]
Arrhythmias	SVM	96.3	First 4 PCA components over action potential and calcium transient biomarkers	86	[290]
Myocardial infarction	Extra trees	80	Gene expression + structural information	251	[291]
Cardiomyopathy	ElasticNet logistic regression	100	Gene expression	31	[292]

3.4.2 Challenges of applications of machine learning in biomedicine

One of the major challenges of machine learning algorithms is to build a model that would robustly perform on the unseen data. Addressing class imbalance is crucial as available datasets are usually heavily unbalanced (i.e.,

unequal distribution between drug types and/or cardiotoxic classes) [293], especially in large datasets that are automatically curated. This limits the generalisation ability of data-driven methods, and in particular of unsupervised ones. Therefore, robust metrics to evaluate the performance of the model are crucial to ensure its predictive power and the translational potential of the model. Generally, data is randomly split into the training set and testing sets, but good practice is also to train and test algorithms with cross-validation and therefore to divide the training set between the actual training set and validation set. Then, the best-performing model is selected based on the metrics calculated for the validation set. The accuracy of the model reflects the number of correct predictions to a total number of predictions. However, accuracy is not a robust metric for a problem with unbalanced classes, which is usually the case with real-world data in predictions of drug cardiotoxicity. Therefore different metrics should be used to account for the specificity and sensitivity of classification models, such as AUC, for example.

Robust cross-validation strategy is yet another challenge. As the aim of cross-validation is to evaluate the ability of the predictive algorithm to extrapolate on a new previously unseen data, the selection of the validation sets is a key step. Random cross-validation, where samples are assigned to actual training and validation sets, are among the common strategies for drug-classification models [289,294]. Yet, if the purpose of a study to build a general approach that would generalise on unseen drugs, models should be validated on the unseen set of drugs and an appropriate cross-validation strategy should be selected. Differences in predictive accuracies between random cross-validation and stratified by subjects were explored in detail by Saeb et al. in a clinical record classification task [295]. The authors clearly demonstrated that algorithms trained on randomly split sets inflate the predictive power of the algorithm and exhibit a significant drop in the accuracy of predicting features of clinical records of unseen subjects from an independent testing set.

3.5 Cardiac human cellular electrophysiological models

Since the first publication of the cardiac model proposed in 1962 by Professor Denis Noble [296], significant progress has been made in the field of cardiac electrophysiology modelling, with new detailed models developed from single ion channel up to whole organ models [297]. As an example, in the last 10 years, a number of human ventricular cell models have been proposed. The general approach is to model protein functions through a series of ordinary differential equations to quantitatively characterise cellular electrophysiology. Among them, the O'Hara-Virag-Varro-Rudy (ORd) human ventricular cellular model [298], published in 2011, was constructed using substantial human ventricular experimental data. A new human ventricular model, ToR-ORd has recently been presented, keeping the strengths and overcoming some of the limitations of the ORd model[299].

At the same time, human iPSCs are gaining increased popularity and have been proved to be a powerful instrument in cardiovascular research, including drug safety testing and disease modelling [300,301]. Following this progress, the first computer model of human iPSC cardiomyocytes, Paci2013, was developed and published by Paci et al. in 2013 [302]. This model has been used since then to predict drug action [303,304] and to model disease phenotypes [305].

Recently the Paci2013 model was updated [306] with a new formulation of Ca^{2+} release in the sarcoplasmic reticulum, replacing the inherited formulation from the TenTusscher2004 model of human adult ventricular cardiomyocytes [307] with a formulation inspired by the Koivumäki2011 model of human adult atrial cardiomyocytes [308]. The resulting model, Paci2018, was calibrated with Ca^{2+} transients and action potential *in vitro* data and extensively validated with *in vitro* data on drug action and on pathological conditions such as hyperkalemia and hypocalcemia. This is the main computational model of cellular electrophysiology that I will use in this thesis in my investigations of hERG trafficking impairment, for correspondence with *in vitro* data collected on human iPSC systems.

3.6 Cardiac cellular electrophysiological models of drug-induced effects

In silico models have shown the ability to account for and predict the clinical risk of drug-induced arrhythmias. This is now a mature field that is being spun out into regulatory guidelines as one of the parts of the CiPA initiative (<http://cipaproject.org/>). Acute drug effects can be simulated at the cellular level by combining single-cell models of cardiac electrophysiology with simple pore-block models based on measures of the potency of drug-channel interactions [290,309,310], or more complex dynamic interactions with some channels [311,312].

Acute drug effects at the cellular level can be simulated by a combination of single-cell models of cardiac electrophysiology and simple pore-block drug models, the latter informed by measures indicating the potency of drug-channel interactions, as described by the Hill equation [290,309,310,313]. For example, this is given for the hERG current as:

$$(3.11) \ g_{Kr} = \bar{g}_{Kr} \left[1 + \left(\frac{[C_d]}{[IC_{50}]} \right)^h \right]^{-1},$$

where $\bar{g}_{Kr}$ is the baseline channel conductance, h is the Hill coefficient, and $[C_d]$ and $[IC_{50}]$ are the tested drug concentration and its concentration for 50% channel inhibition, respectively.

Another approach is to simulate such drug-cell interactions using more complex dynamic models, which has also been explored for hERG [311,312], and is illustrated in Figure 3-1 for the Markov model of hERG proposed as part of the CiPA initiative.

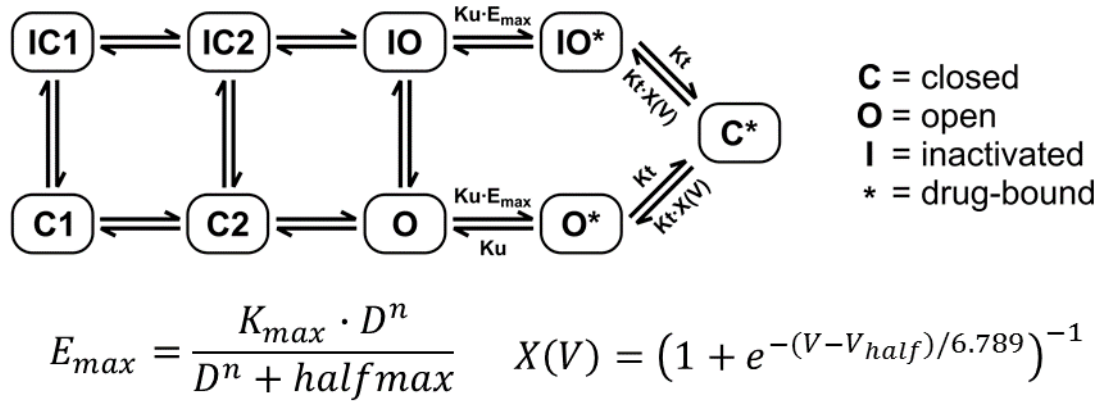


Figure 3-1 Markov model of the hERG channel for acute drug-channel interactions (Li et al., 2018). Model parameters are obtained from in vitro experiments at physiological temperature. Ku: drug unbinding rate; Kmax: maximum drug effect; n: Hill coefficient of drug binding; halfmax: power of the half-maximal drug concentration; Vhalf: drug trapping potential. Adapted from https://github.com/FDA/CiPA/tree/master/hERG_fitting.

The aforementioned approaches reproduce experimental data and have proved useful in simulating disease conditions or variability of acute drug responses within healthy or diseased populations. They also allow modelling of multi-channel pharmacology by introducing the drug-induced changes in separate currents properties.

3.6.1 Biomarkers for quantifying cardiac electrophysiological model output

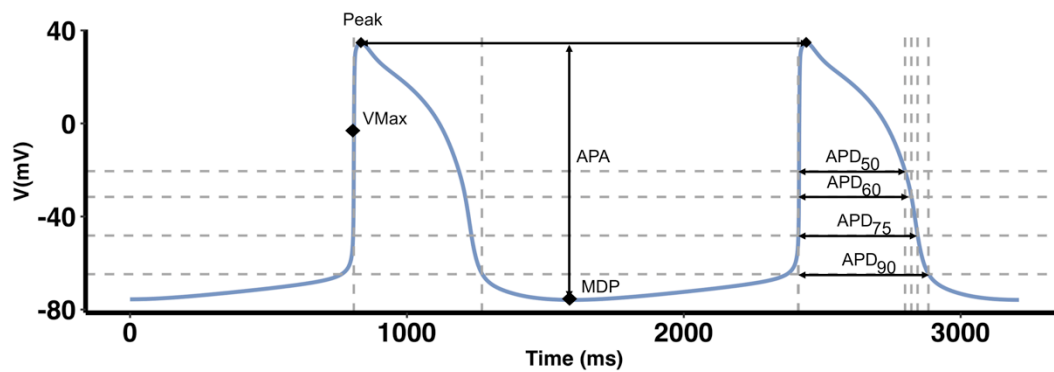


Figure 3-2 Action potential biomarkers calculated in this thesis. Peak: peak membrane potential; APA: action potential amplitude; VMax: maximum upstroke velocity; MDP: minimum diastolic potential; APD: action potential duration.

Biomarkers are characteristics that describe the output of electrophysiological models, including action potential features. In Figure 3-2, I have described the biomarkers used in this thesis, standard in previous studies. These include the rate of spontaneous action potentials (Rate), minimum diastolic potential (MDP), peak membrane potential (Peak), action potential amplitude (APA), maximum upstroke velocity (VMax), action potential duration (APD). More specifically, MDP is the most negative membrane potential attained during the cardiac cycle, Peak is the peak value of action potential, and APA is measured as a difference between Peak and MDP. From these values, APD can be measured at different percentages of repolarisation. VMax is the maximum value of the change in membrane potential with respect to time during the upstroke of an action potential.

3.7 Databases for drug discovery

A decade of rapid data accumulation in the public domain has resulted in multiple data repositories, providing scientists with thousands of high-quality samples. The growth of such databases has stimulated drug discovery. In the following section, I will briefly review pathway databases, gene expression databases and side effect databases.

3.7.1 Pathway databases

There are two pathway databases freely available online: Kyoto Encyclopedia of Genes and Genomes (KEGG) [314] and Reactome [315]. The KEGG pathway database contains information on 512 pathways, with topology and cell localisation for each pathway element. At present, Reactome is the most comprehensive pathway database and contains information on over 10,000 genes assigned to over 2000 pathways structured as gene sets, but it provides no information on the pathway topology. Importantly, according to the latest estimations, the human genome contains around 19,000 to 20,000 protein-coding genes [316], so even the most comprehensive Reactome annotation covers only half of the protein-coding transcriptome. In addition, pathway databases such as KEGG include limited disease states and mostly

describe cancer-related pathways. Therefore, pathway analysis tools, such as GSEA, SPIA and iPANDA, should be used with caution considering their limitations as supervised methods.

3.7.2 Gene expression databases

Launched in 2014, the NIH-supported Library of Integrated Network-Based Cellular Signatures (LINCS) project is the largest catalogue of cellular responses (cellular signatures) on small molecules. LINCS contains gene expression profiles, fluorescent imaging profiles, etc. LINCS data were generated by several data centres. The LINCS Center for Transcriptomics and the Drug Toxicity Signature (DToxS) Generation Center were responsible for the generation of gene expression profiles of small molecules incubated with human cell lines. The efforts of the LINCS Center for Transcriptomics resulted in the Connectivity Map database [317], which consists of over 1.6 million samples profiled for over 8000 small molecules. Such a scale became possible thanks to the development of the L1000 microarray assay, which captures expression levels for about 1000 of the most representative landmark genes. The database also includes different concentrations of small molecules incubated for 6, 12, 24 or 48 hours with cancer cell lines that have already demonstrated clinical relevance [318], and mRNA measured from samples incubated with small molecules and for matching control samples (vehicle without a small molecule). The DToxS Center database contains information on 53 drugs incubated with up to four induced cardiomyocyte lines. mRNA levels were measured using RNAseq technology, which is able to capture the whole transcriptome (more than 20,000 genes). Although still growing, the DToxS database is limited by the number of drugs compared to the comprehensive Connectivity Map database.

3.7.3 Side effect databases

The pathway and gene expression databases mentioned above could be linked to various drug identifiers, such as Broad ID or PubChem CID, as well as metadata, such as drug targets and, in particular, drug side effects.

Two of the largest databases of recorded, marketed adverse drug reactions that are publicly available are the FDA Adverse Event Report System (FAERS) and the Side Effect Resource (SIDER) [319], which contains processed FAERS data. Observed side effects are reported by healthcare professionals, so the terminology is not unified. For example, according to the Medical Subject Headings (MeSH) classification “cardiotoxicity is damage to the heart or its function secondary to exposure to toxic substances such as drugs used in CHEMOTHERAPY; IMMUNOTHERAPY; or RADIATION”, while the National Cancer Institute (NCI) defines cardiotoxicity more generally as “toxicity that affects the heart”. At the same time, the current International Classification of Diseases system, now in its 10th revision (ICD-10), has no medical code for cardiotoxicity. It should be clearly emphasised that FAERS side effect records do not indicate the causal relationship between a drug and its reported side effects, but the *potential safety issue*. SIDER contains machine-processed FAERS information on adverse reactions, which is filtered by semantic type with manually selected side effect terms. Therefore, SIDER is better structured and includes refined data; however, it is not protected from ambiguous terminology. These peculiar properties should be taken into account in data collection along with the bias in the information, that is based on the reports from American studies and practice, as they were collected using the FDA reporting system.

3.8 Conclusions

In this chapter, I began by outlining the previous literature on transcription-based biomarkers of drug-induced cardiotoxicity and machine-learning methods for predicting cardiac side effects. I also highlighted the key challenges of machine-learning models in drug classification: unbalanced data and the need for robust performance measures and adequate cross-validation strategies. The aims of this chapter are to motivate a transcriptional analysis coupled with machine learning to predict the mode of drug actions and adverse cardiac reactions, and to introduce the modelling of drug actions on electrophysiology along with the biomarkers used to describe those models.

While the approaches described in this chapter have proved to be extremely useful, the chronic effects of drugs are less explored. The complexity of drug interaction with cells and tissues described in Chapter 2 suggests that suitable models for capturing these interactions are required in order to augment the predictive value of current in silico approaches.

The focus of the next chapter, Chapter 4, is to present the analysis of the comprehensive drug dataset collected and the first results of this thesis. I start by describing the underlying structure of the data, including the distribution of safe drugs known to cause pharmacological complications. At the same time, I observed dependencies between different drug-induced cardiac disorders. The results of this analysis motivated me to use the chain of classifiers approach described in Chapter 5. Finally, in Chapter 6 I propose a mechanistic model of hERG channel trafficking, and in Chapter 7 I describe the results of drug-induced channel inhibition modelling.

4 Prevalence of cardiotoxicity forms: data analysis

This chapter presents the work derived from the publication: “Dual transcriptomic and molecular machine learning predicts all major clinical forms of drug cardiotoxicity”

4.1 Introduction

At the end of Chapter 3, I described how gene expression signatures could be used as a biomarker of cell response to cardiac dysfunction and drug action. This led to my interest in collecting a large database of drugs that are known to cause cardiotoxicity in clinic, and drugs that have no reports of cardiac toxicity. From this point I explored large publically available knowledge and data repositories.

In this chapter, I describe the data collection and analysis of existing association between drug-induced cardiac disorders and their prevalence within different pharmacological classes of drugs.

4.2 Methods and Materials

4.2.1 Data collection

I used the following databases and knowledge repositories to collect the dataset:

- DrugBank ([320]; www.drugbank.com): a database of drug information, including drug targets;
- Connectivity map Project ([317]; <https://clue.io/cmap>): a catalog of cellular responses (cellular signatures) on chemical (drugs) and genetic perturbations;
- SIDER ([319]; <http://sideeffects.embl.de/>): a database of observed side effects of drugs reported by healthcare professionals;

- MedDRA (<https://bioportal.bioontology.org/ontologies/MEDDRA>): a medical terminology dictionary that is also used by regulatory agencies to describe drug safety information;
- MESH (<https://www.ncbi.nlm.nih.gov/mesh>): a control medical vocabulary that provides clinical information about drugs including their therapeutic classes.

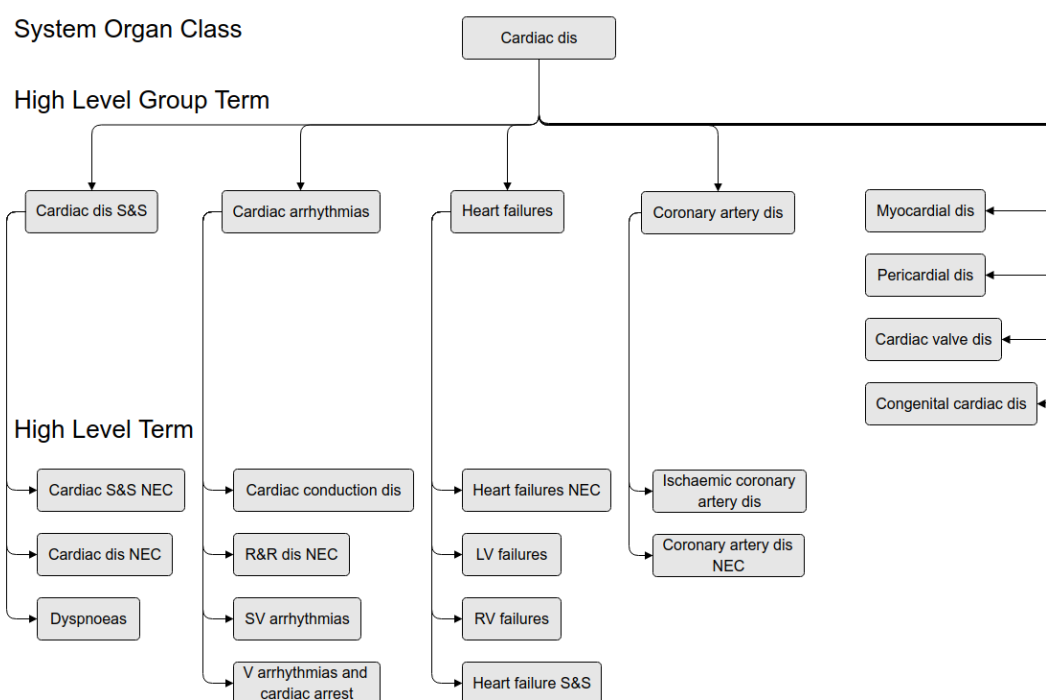


Figure 4-1 MedDRA classification for cardiac disorders in the SIDER database. Each High Level Term is linked to only one High Level Group Term. For example, 'Cardiac disorders' belongs to System Organ Class and unites 8 High Level Group Terms. Dis: disorders; S&S: sign and symptoms; NEC: not elsewhere classified; R&R: rate and rhythm; SV: supraventricular arrhythmia; V: ventricular; LV: left ventricular; RV: right ventricular.

Names and IDs were retrieved from the MedDRA dictionary for all cardiac (MedDRA ID 10007541; Figure 4-1) and vascular (MedDRA ID 10047065) disorders. Interactions of chemicals in the form of STITCH compound identifiers and their MedDRA terms of side effects were downloaded from the publicly available the SIDER database.

I used PubChem Compound Identifiers (CIDs) to match this list to the information on drug targets, drug status ('approved', 'investigational', etc.),

and drug SMILES notations obtained from the Drugbank database. Only experimentally verified interactions between drugs and proteins that are targeted by those drugs (as provided by Drugbank) were considered for this work. Compounds linked to the MedDRA term 'cardiac disorders' were labelled as drugs with cardiotoxicity reports. Compounds with same targets and no record of cardiac disorders in the database were considered as safe. Safe compounds were additionally filtered by their status, and only approved compounds currently on the market were used for analysis, resulting in 26 (out of 759) drugs being removed from the analysis. This was performed to prevent possible unreported toxicities in drugs considered safe but withdrawn from the market. The unsafe group includes both marketed and withdrawn drugs, including drugs withdrawn due to cardiac side effects.

In this work, I used the Connectivity map project as a source of gene expression cell responses to drugs, or drug transcriptional profiles. These were measured using an L1000 high-throughput profiling method. The L1000 fluorescent assay allows the detection and quantification of the expression of up to 978 landmark and 80 control transcripts simultaneously in each well of 384-well plates, where each well can contain a separate drug profile. This massive scale expression data is available in multiple levels starting from raw fluorescent intensity values from each well to replicate collapsed scores for drugs. In this work, I explored the provided transcriptional profiles (normalised across each scan plate) of the 977 'landmark genes' (Level 3a – NORM, as described in detail in https://clue.io/connectopedia/data_levels). I used transcriptional profiles of six cell lines (A549 and MCF7, both with well-established culture and treatment protocols, and with most of the data profiled, and PHH, SKB, SKM1, A673 used for sensitivity testing), two incubation times (6 and 12 hours), and multiple drug concentrations. To link the drug transcriptional profiles provided by the Connectivity map project to the information about their side effects, targets and status, I utilised the Chemical Translation Service (<http://cts.fiehnlab.ucdavis.edu/>) to match PubChem CIDs to their corresponding Broad IDs. I used MESH vocabulary to extract information about therapeutic classes of drugs. The following drug classes were analysed: Anti-allergic (7 drugs), Anti-inflammatory (83 drugs),

Antimetabolites (53 drugs), Antineoplastic (93 drugs), Antiviral (47 drugs), Central Nervous System (538 drugs), Cardiovascular (160 drugs), Dermatological (6 drugs), Gastro (8 drugs), Hormonal (85 drugs), Immune (5 drugs), Metabolic (1 drug), Metabolites (23 drugs), Respiratory (1 drug), Skeletal (9 drugs), Vitamins (13) and Other (1 drug).

In total, I collected a database of 1,131 drugs (fully analysed for relationships across cardiotoxicity types), 357 of which had transcriptional profiles (used for prediction), with a total of 9,933 samples. Samples refer to independent transcriptional profiles of a drug at a given cell line, incubation time or concentration.

4.2.2 Calculation of molecular descriptors and fingerprints

I used the 'rcdk' package [321] to calculate seven molecular descriptors widely used in drug property prediction [322]: molecular weight (MW), partition coefficient (XLogP), atomic polarisabilities (apol), topological polar surface area (TopoPSA), polar surface area expressed as a ratio to molecular size (tpsaEfficiency), Ghose-Crippen LogKow (ALogP) and molar refractivity (AMR). I also calculated the commonly used 79 bit 'estate' fingerprint.

4.2.3 Training testing set design

I randomly split the entire drug dataset by protein targets into unique training (291 drugs, 8,237 samples) and testing (66 drugs, 1,696 samples) sets. By such design, both datasets only overlap by a set of protein targets, but drugs on the testing phase are completely unseen during training, therefore facilitating preclinical translation to novel chemicals. These training and test datasets will be used in the development of the machine learning algorithms for the prediction of drug cardiotoxicity presented in next chapter, and explored here to visualise their data structure. I collapsed samples, so each drug profile referred to gene expression values for one cell line, incubation

time and concentration. I used this selected trained matrix to first select the relevant transcriptional features size of 340×1154 and tested on 340×746.

4.2.4 Feature selection methods

I applied two selection methods (correlation-based, wrapped-based) to identify the relevant to cardiotoxicity genes. I applied feature selection to the training set to ensure that the testing set was independent.

Correlation-based feature selection allows identification of features that highly correlated with the target label, or cardiotoxicity type in this case, and uncorrelated to each other to reduce redundancy. As a correlation-based method, the 'select.cfs' function from the Biocomb R package [323] was used to select the list of genes by their individual predictive ability for each label.

Wrapper-based methods evaluate subsets of features in an iterative manner to find the features that are mostly predictive of a target label (cardiotoxicity form). Unlike correlation feature selection, it allows to detect possible interactions between features. As a wrapper method, the Boruta algorithm implemented in the Boruta R package [324] was used to select the list of relevant features through their importance in predicting each label.

After I selected the subset of genes for each cardiotoxicity type using those two methods, I compared them using the measures described below.

4.2.5 Measure of association

To analyse the agreement between different cardiotoxicity forms I calculated Cohen's Kappa scores [325], using the 'Kappa.test' function from the fmsb R package:

$$(4.1) \text{Kappa} = 1 - \frac{1-p_o}{1-p_e},$$

where p_o is the relative observed agreement between two binary vectors, and p_e is the expected agreement between predicted and actual values. Values smaller than 0 demonstrate poor agreement and values from 0.81 to 1 correspond to almost perfect agreement. These scores are used in this chapter to evaluate similarities, first between the binary vectors of cardiotoxicity types, and second between vectors of transcriptional features selected for those disorders. They are also used in next chapter to estimate the accuracy of classifiers.

4.2.6 Visualisation of hidden patterns in the data

I used t-distributed stochastic neighbour embeddings (tSNE) to visualise the structure of the dataset and features selected. The tSNE algorithm calculates non-linear representations of the data. First, the algorithm computes conditional probabilities between drugs samples proportional to the similarity of those samples in a high-dimensional space. After, it defines the Student t-distribution to recreate this probability distribution on a low-dimensional space.

I used the 'cude.tsne' R package to calculate tSNE representations in 2-D space. I explored three values for the number of drug samples that were considered as neighbours and so set to be similar (perplexity values): 10, 30 and 50.

4.3 Results

4.3.1 Frequency analysis of cardiotoxicity forms

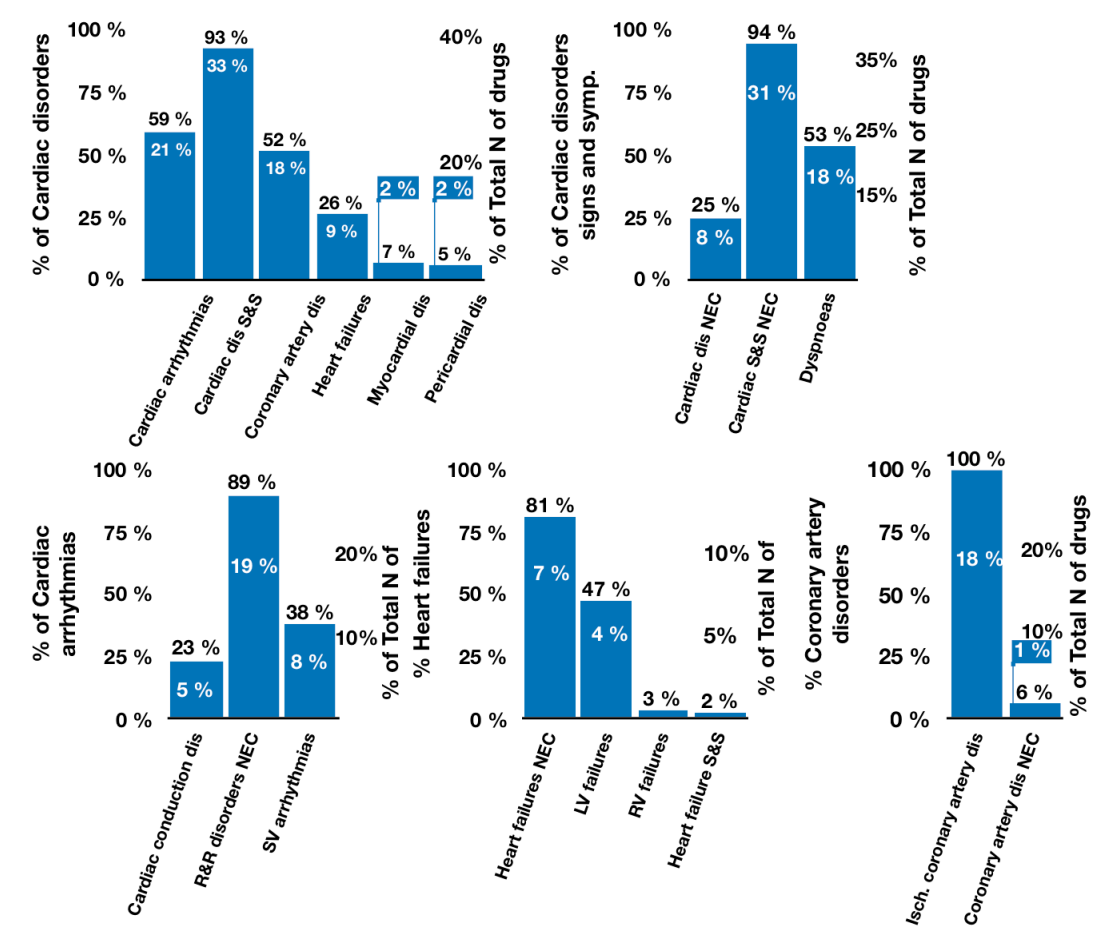


Figure 4-2 Proportion of drugs labelled as unsafe for given cardiotoxicity forms of High Group Level and High Level Terms, out of all unsafe drugs (numbers in black) and all drugs (numbers in white). Dis: disorders; S&S: sign and symptoms; NEC: not elsewhere classified; R&R: rate and rhythm; SV: supraventricular arrhythmia; V: ventricular; LV: left ventricular; RV: right ventricular;

To explore the relation between transcriptional profiles and structure of drugs with their clinical cardiac safety, I first collected a large dataset of medications linked to their information about presence or absence of reports on cardiotoxic adverse events, their protein targets and enzymes, their structure, and their gene expression signatures from treatment on a variety cells.

In total, I collected information about 1,131 drugs, 35% of which with reported cardiotoxicities (Figure 4-2). 'Cardiac disorders signs and symptoms' is the

most common cardiotoxicity form reported, with 371 out of 1,131 drugs labelled as unsafe, followed by ‘cardiac arrhythmias’ (237 drugs), ‘coronary artery disorders’ (206 drugs), ‘heart failures’ (104 drugs), ‘myocardial disorders’ (26 drugs) and ‘pericardial disorders’ (21 drugs). An uneven distribution of drugs with and without cardiotoxicity reports was also observed for all term levels (Figure 4-2).

4.3.2 Analysis of drug-induced cardiac disorders’ association

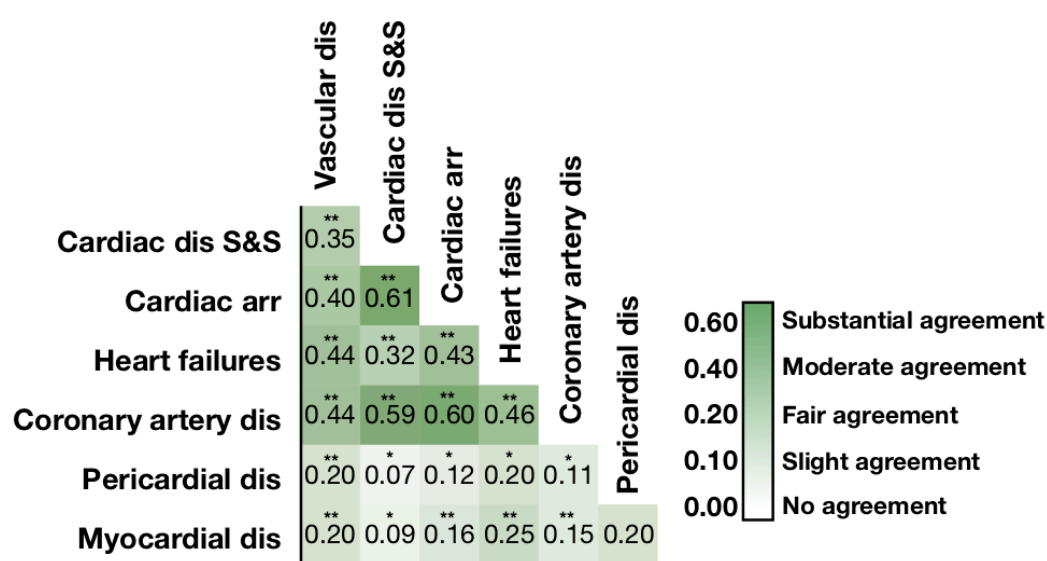


Figure 4-3 Levels of association between cardiotoxicity forms, measured in terms of Cohen's Kappa. Symbols represent statistical significance level by Z-test (* $p < 0.05$, ** $p < 0.01$). ‘Cardiac dis S&S’: MedDRA term for cardiac disorder signs and symptoms; ‘dis’: for disorders.

Figure 4-3 illustrates the association between cardiotoxicity forms in terms of Cohen's Kappa. I found that ‘Cardiac arrhythmias’ demonstrate a substantial association with both ‘cardiac disorder signs and symptoms’ and ‘coronary artery disorders’, with Cohen's Kappa values of 0.61 and 0.60, respectively. Interestingly, ‘heart failure’ demonstrates a lower agreement with ‘myocardial disorders’ (0.25) and ‘cardiac disorder signs and symptoms’ (0.32), compared to ‘cardiac arrhythmias’ (0.43) and ‘coronary artery disorders’ (0.46).

To obtain the list of transcriptional features that are the most predictive of the six major forms of cardiotoxicity, I applied two feature selection methods:

correlation-based and a wrapper-based method. The first one identified 125 genes as important for all six cardiotoxicity forms, while 159 genes were identified by the second. Interestingly, only 31 of those genes were common. I then mapped those genes to Reactome pathways, resulting in improved agreement between both methods. For example, the two algorithms did not identify common genes for heart failure or coronary artery disorders, whereas they shared common Reactome pathways terms such as ‘G-alpha (i) signalling events’ for heart failure, and ‘Ub-specific processing proteases’ and ‘Regulation of TP53 Degradation’ for coronary artery disorders. Interestingly, G protein-coupled receptor transduction was selected as important by both methods for the prediction of cardiac disorder signs and symptoms (‘G alpha (q) signalling events’ and ‘G alpha (s) signalling events’) and heart failure (‘G-alpha (i) signalling events’). IGF1R and IGF1R-related signalling were also among the genes and Reactome terms selected by both methods for cardiac disorder signs and symptoms and for pericardial disorders. MAMLD1 gene, included in the Notch signalling pathways, was the only one ranked as important for predicting cardiac arrhythmias by both selection procedures.

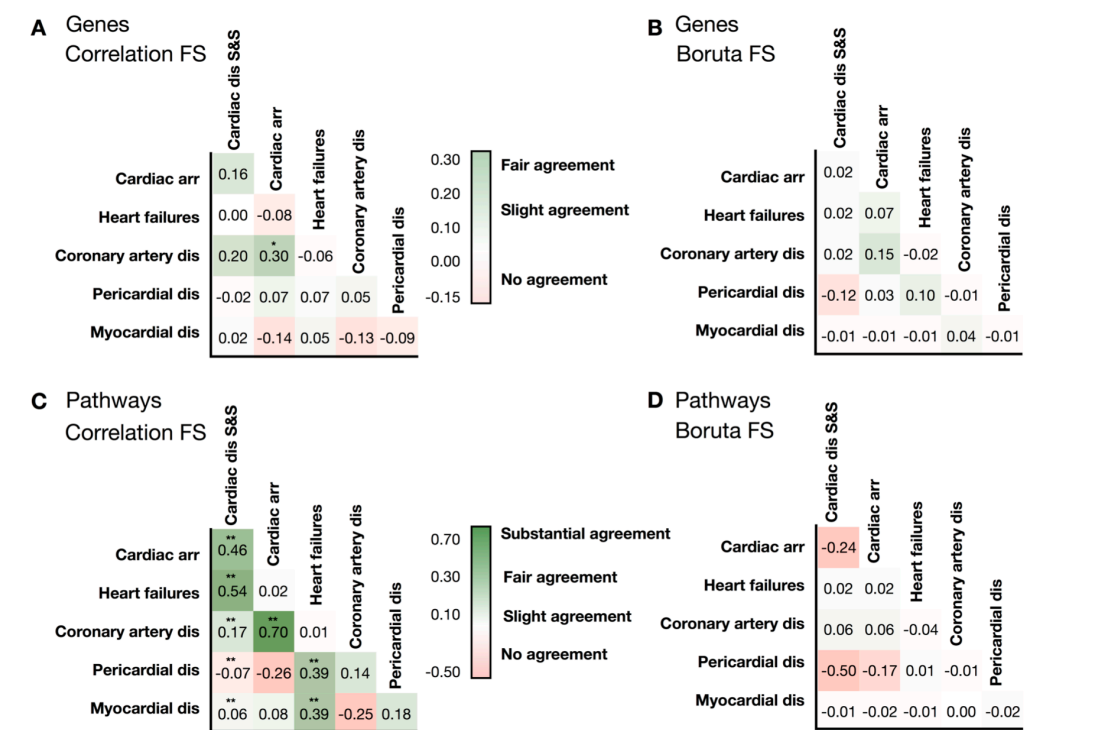


Figure 4-4 Similarities in the list of selected genes identified by (A) the correlation feature selection, and (B) the Boruta wrapper-based algorithm. Similarities in the list of pathways associated with genes identified by correlation-based feature selection (C), and the wrapper-

based algorithm (D). Levels of association between cardiotoxicity forms are measured in terms of Cohen's Kappa. Symbols represent statistical significance level by Z-test (* $p < 0.05$, ** $p < 0.01$). 'Cardiac dis S&S': MedDRA term for cardiac disorder signs and symptoms; 'dis': for disorders.

Figure 4-4 shows the association between cardiotoxicity forms for a selected set of transcriptional features. I hypothesized that transcriptional data would reveal underlying information on cardiotoxicity types when genetic pathways, rather than individual genes, are considered in the analysis. Figure 4-4 A and B show the comparison of the gene vectors ranked as important by two feature selection methods for the different cardiotoxicity forms. Similarly, Figure 4-4 C and D show this comparison in terms of the Reactome pathways to which those genes are related. It can be seen that 'coronary artery disorders' also shows a high level of association with 'cardiac arrhythmias' on gene and pathway level for correlation-based selected by the correlation-based method. I also found that there is a low agreement between feature sets selected by the Boruta method.

4.3.3 Visualisation of the dataset structure

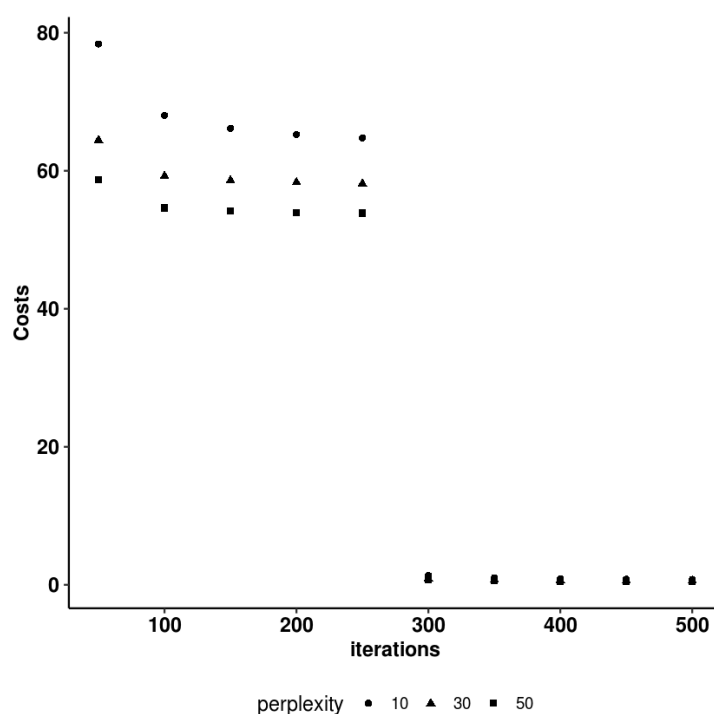


Figure 4-5 Changes in similarity (cost) between a cluster of drug samples for varying perplexity (number of neighbouring samples considered to be close to each other) by tSNE.

After selecting the feature set predictive of all major cardiotoxicity types, I decided to explore the underlying structure in data related to cardiotoxicity types or possible biases. I first explored how a different number of neighbouring samples considered similar would affect the measurement of the similarity between clusters on the data (KL-divergence), by varying the perplexity parameter of the tSNE algorithm.

Figure 4-5 shows that there is no prominent difference between costs, and that the algorithm needs 300 iterations to converge and reach values close to 0 for each value of perplexity.

I then explored how perplexity values would affect the visually observed clustering of drug samples. As shown in Figure 4-6, for larger perplexities (values of 30 and 50), samples are rather clustered by cell lines, incubation times and dose (with smaller doses forming small clusters), but not by pharmacological classes. For the perplexity value of 10, clusters cannot be fully explained by cell lines, incubation times and dose or pharmacological classes. Therefore, I selected the perplexity value of 10 to explore if those clusters can be partially explained by the cardiac safety of drugs.

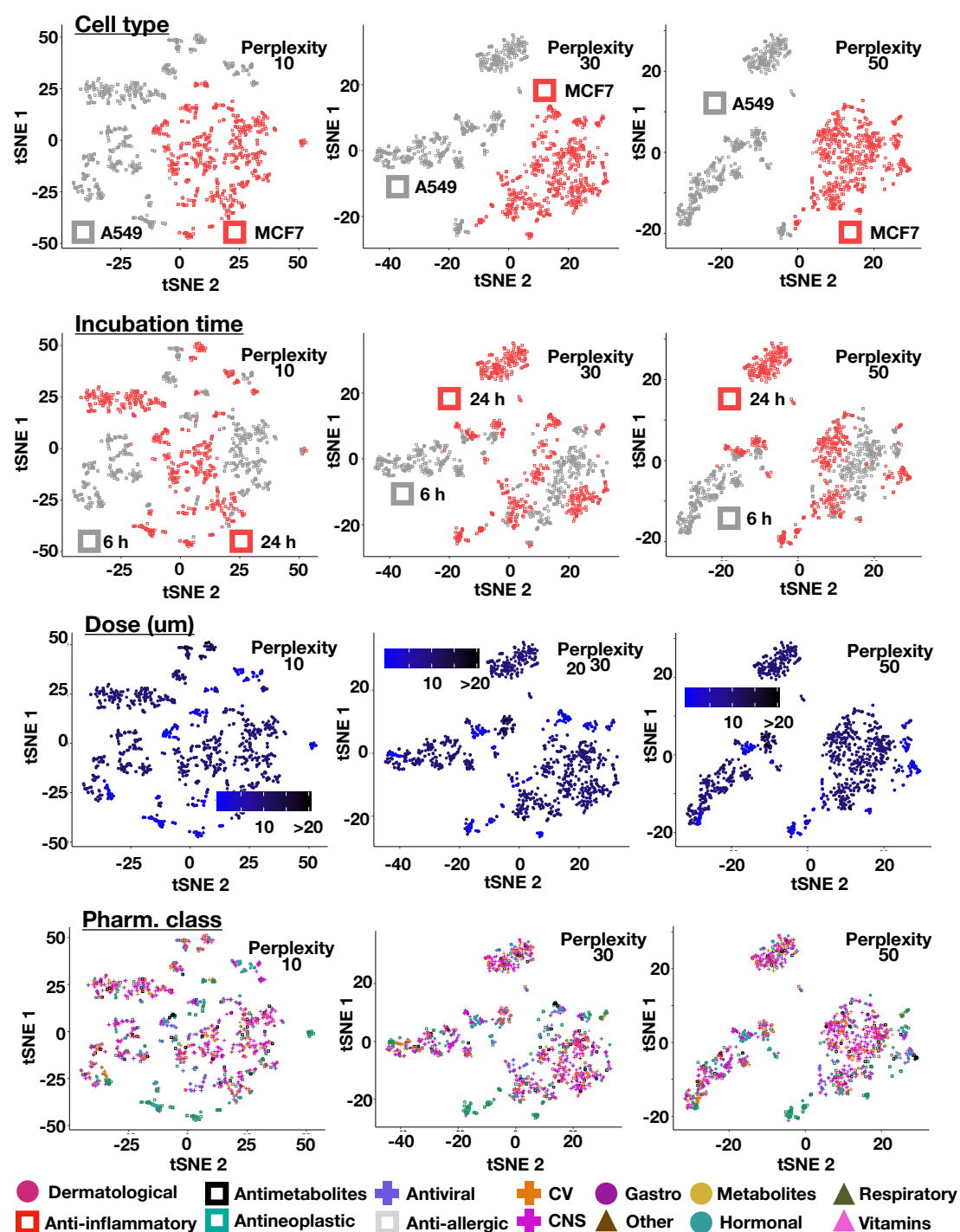


Figure 4-6 Learnt 2D representations of the data show clustering by cell type, dose and incubation time. tSNE mapping of 340 selected transcriptional features and molecular descriptors and fingerprints of 1154 drug samples in the training set to 2D space, for varying perplexity values, and coloured by cell types (A549 and MCF7), incubation times (6 and 24 hours), dose and pharmacological classes (14 classes).

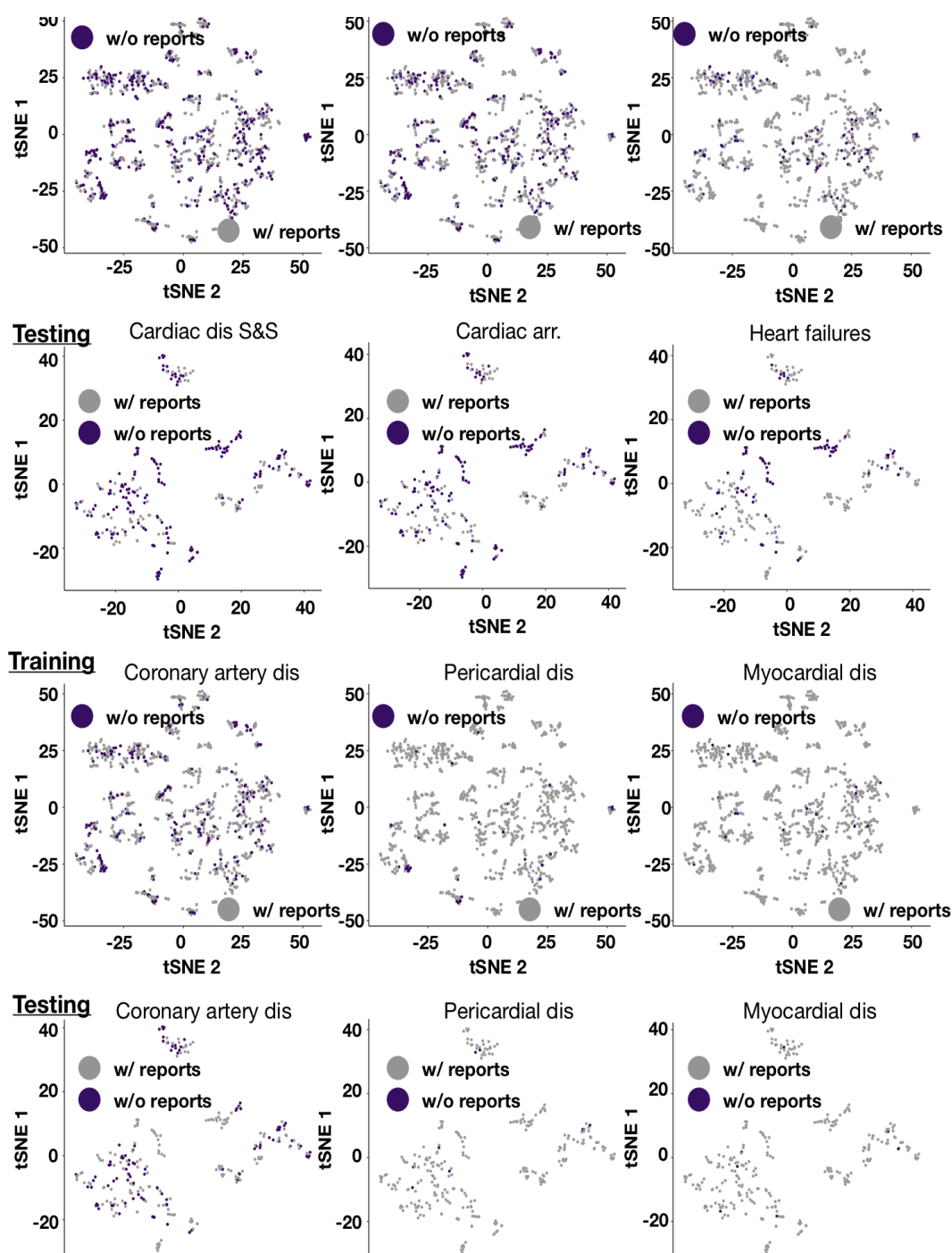


Figure 4-7 Learnt 2D representations of the data show no clustering by cardiotoxicity forms. 2D tSNE mappings of 340 selected transcriptional features, and molecular descriptors and fingerprints, of 1154 drug samples in the training set and 265 samples in the testing set (perplexity 10). Samples are coloured by the presence (w/ reports) or absence (w/o reports) of the adverse reports for six cardiotoxicity forms: Cardiac disorders signs and symptoms, cardiac arrhythmias, heart failures, coronary artery disorders, pericardial disorders and myocardial disorders. 'Cardiac dis S&S': MedDRA term for cardiac disorder signs and symptoms; 'dis': disorders.

Figure 4-7 shows the learnt t-SNE representation in low-dimensional space (2-D) for each six forms of cardiotoxicity. Drug samples show no clustering according to cardiotoxicity types either in testing or training sets, suggesting complex non-linear dependencies between transcriptional features, molecular descriptors and fingerprints, and cardiac safety. A similar absence of clustering by drug safety was observed for vascular disorders (Figure 4-8).

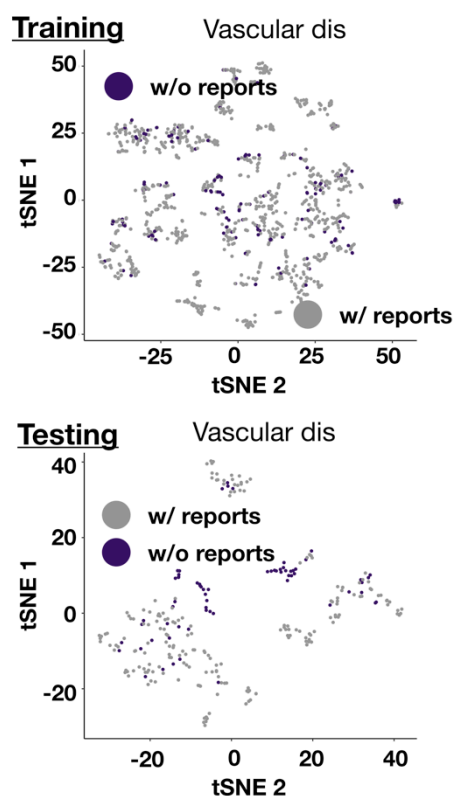


Figure 4-8 Learnt 2D representations of the data show no clustering for vascular disorders. 2D tSNE mapping of 340 selected transcriptional features, and molecular descriptors and fingerprints, of 1154 drug samples in the training set and 265 samples in the testing set. Samples are coloured by presence (w/ reports) or absence (w/o reports) of adverse reports for vascular disorders. 'dis': disorders.

4.3.4 Prevalence of specific forms of cardiotoxicity associated with drug pharmacological classes

Table 4-1 Number of drugs with and without reported six forms of cardiotoxicities for four major pharmacological classes. CNS: central nervous system; CV: cardiovascular; S&S: signs and symptoms; dis: disorders.

Pharmacological class	Reports of cardiac toxicity	Cardiac dis S&S	Cardiac arrhythmias	Coronary artery dis	Heart failures	Myocardial dis	Pericardial dis
CNS	yes	164	129	102	49	16	9
	no	374	409	436	489	522	529
CV	yes	61	34	36	24	2	1
	no	99	126	124	136	158	159
Anti-inflammatory	yes	24	11	10	4	2	1
	no	59	72	73	79	81	82
Antineoplastic	yes	46	25	23	13	3	7
	no	47	68	70	80	90	86

I finally analysed the prevalence of specific cardiotoxicity forms across drug classes. Table 4-1 shows the number of drugs with and without reports of 'cardiac disorders signs and symptoms', 'cardiac arrhythmias', 'coronary artery disorders', 'heart failures', 'myocardial disorders', 'pericardial disorders', for four major pharmacological classes. For example, I found that 49% (46 out of 93) of antineoplastic drugs are reported to cause 'cardiac disorders and signs and symptoms', indicating a high prevalence across cardiotoxicity forms and drug classes. Equivalently, 24% of CNS, 21% of CV, and 27% of antineoplastic agents produced cardiac arrhythmias, and 23% of CV and 25% of antineoplastic agents induced coronary artery disease. On the other hand, the prevalence of heart failure, myocardial disorders and pericardial disorders was lower for all drug classes, although still significant in some cases (for example, 14% of antineoplastic agents producing heart failure).

4.4 Discussion

In this study, I collected a large database of drugs and I linked them to their cardiac and vascular safety information. The collected database contains information of 1,131 drugs, 35% of which are known to induce at least one out of six types of cardiotoxicity: 'Cardiac disorder signs and symptoms', 'Cardiac arrhythmias', 'Heart failure', 'Coronary artery diseases', 'Pericardial disorders' and 'Myocardial disorders'.

The collected dataset shows an imbalance in terms of the number of drugs with and without cardiac adverse reaction reports, which is in line with known challenges in biomedical data analysis [326]. At the same time, the observed prevalence in cardiotoxicity among pharmacological classes is in agreement with the literature (see Chapter 2), with antineoplastic and central nervous system agents as groups causing most of the known adverse events.

Previous studies have established connections between cardiovascular disease and other chronic conditions, which is considered important because it poses a greater risk for patients [327]. Similarly, further research on associations between drug-side effects can provide additional insights in cardiac safety, and can help to inform hypotheses of drug action [328]. In the curated dataset I collected, I also observed associations between forms of drug-induced adverse effects. Clinical reports have evidenced a significant association between heart failure and other cardiovascular comorbidities, such as atrial fibrillation, ischemic heart disease and arrhythmias [327,329]. Patients with a history of coronary heart disorders have been also shown to have a higher incidence of atrial fibrillation, one of the most common forms of cardiac arrhythmias [330]. In this study, coronary artery disorders, which include ischemia and myocardial infarction, demonstrated a significant similarity to cardiac arrhythmias in terms of the drugs they are related to. I showed the same for heart failure and cardiac disorders, which displayed a moderate and substantial agreement to cardiac arrhythmias.

Notably, the list of genes identified as most important features (obtained by using two completely distinct feature selection methods while counting distinct genes for each label) demonstrates a significant amount of intersection at the level of associated pathways. For example, both methods ranked pathways related to Notch signalling as important for cardiac arrhythmia prediction. Previous evidence has shown the importance of Notch signalling in heart development and cardiac disease, including malignant congenital arrhythmias [331]. IGF signalling and IGF1R were also selected consentaneously by both algorithms for cardiac disorder signs and symptoms and pericardial disorders, evidencing their key role in heart tissue functioning [332]. In spite of being profiled in cancer cell lines, the selected features seem biologically relevant to cardiotoxicity, given their human origin. This emphasises the potential benefits of using the proposed combined approach for feature selection.

4.5 Conclusions

In this chapter, I have collected and curated a large comprehensive dataset of 1,131 drugs from publicly available resources. I have explored the frequency and distribution of cardiotoxicity forms, and showed that this distribution is unbalanced between safe drugs and drugs with reported cardiotoxicity. This can represent additional difficulties for machine learning approaches, where balancing of the dataset should be considered, as well as the selection of suitable measures of model performance. Equivalently, cardiotoxicity forms show associations in terms of co-occurrence and the transcriptional features selected as predictive. Therefore, models capable to retain relations between labels should be explored. The visualisation of the underlying data structure showed that drugs with and without reported cardiotoxicity are neither linearly separable nor easily predicted by manifold methods, and therefore non-linear machine learning models should be considered for this task. However, drug samples were clustered by cell types, incubation times and dose, and hence the performance of these approaches should be explored in detail for those characteristics. My resulting study on these topics, which introduces machine learning prediction of six cardiotoxicity forms, is described in the next chapter.

5 Dual machine learning prediction of drug-induced cardiotoxicity

This chapter presents the work derived from the publication: “Dual transcriptomic and molecular machine learning predicts all major clinical forms of drug cardiotoxicity”

5.1 Introduction

In Chapter 4, I described the dataset of 1,131 drugs that I curated and collected from publicly available resources, showing that it is imbalanced and has an uneven distribution of compounds known to cause cardiac adverse effects in human. I also identified associations between cardiotoxicity forms, showing that safe and compounds with reports are not linearly separable.

In this chapter, I address these challenges posed by data in drug cardiotoxicity and present a machine learning approach for the prediction of six drug-induced cardiotoxicity forms. I first will formulate the classification problem, after which I will introduce study design, balancing approach and cross-validation strategies. I will then use the data collected and described in the previous chapter to build different predictors of cardiotoxicity and explore their predictive power across cardiac disorders, drug classes, cell lines and types of features (transcriptomic and molecular descriptors, investigated both independently and in a dual approach, to discriminate the separate and joint predictive power of these data sources).

5.2 Methods

5.2.1 Study design

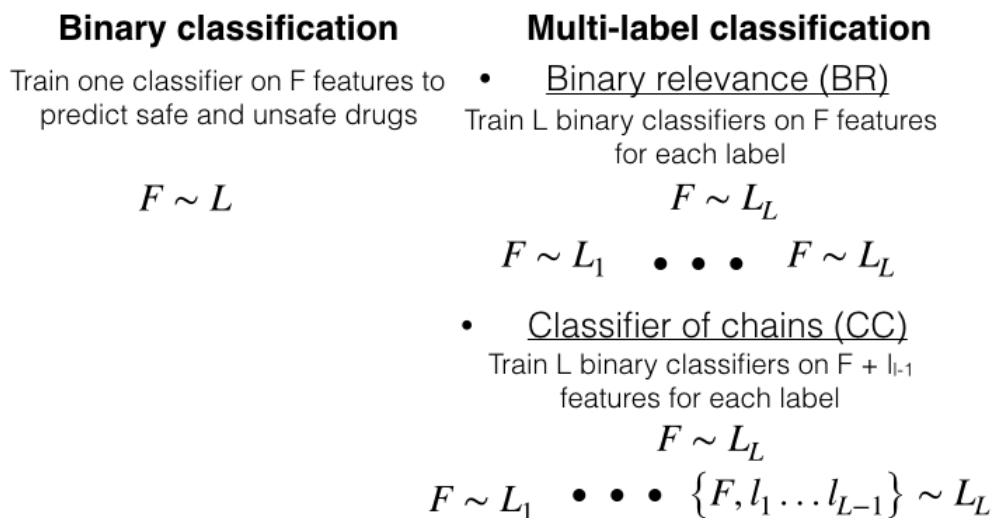


Figure 5-1 Two proposed approaches for classification. In this study, binary relevance classification (a set of independent classifiers – BC) is compared to chain of classifiers (CC).

As described in Chapter 5, the collected dataset contains several forms of cardiotoxicity. For the classification problem, I consider six cardiotoxic forms, where each form categorises drug samples into two classes: with and without known cardiotoxicity. Here, I will explore two approaches for multi-label classification: a set of individual binary classifiers (BC) and a chain of classifiers (CC), as shown in Figure 5-1.

Figure 5-2 describes the computational and data framework proposed in this study. After data collection and isolation of a testing set (completely unseen during training, see “Training testing set design”, Chapter 4, page 85), I selected the set of most descriptive features using the training set. Transcriptional and molecular descriptors for the drugs were used as inputs to the machine learning algorithms.

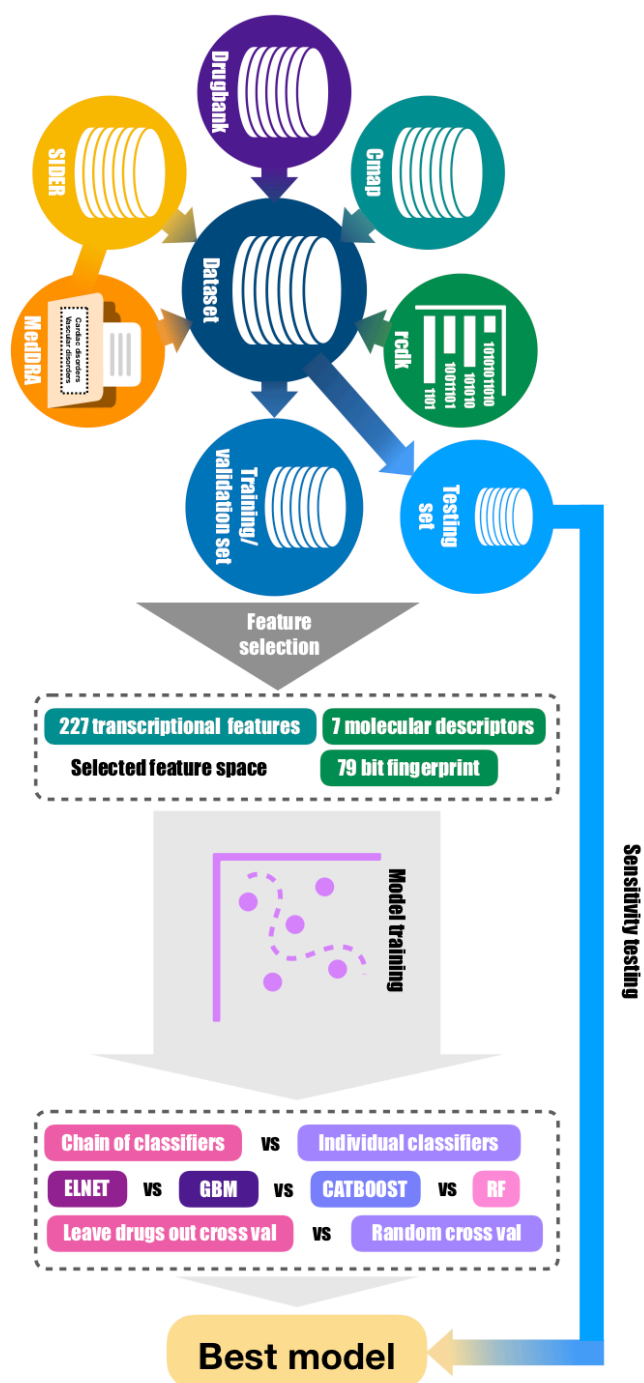


Figure 5-2 The design of the proposed machine learning model builds on a comprehensive database of drugs linked to their transcriptional profiles, molecular descriptors, fingerprints and safety information. After collation, the whole dataset was split into training and testing sets, of unique sets of drugs in each. Using training data, the set of the most predictive features was selected. Those features were later used to train multi-label models (chain of classifiers with nested stacking vs sets of individual classifier). Training was performed with two cross-validation strategies: random vs leave-drug-out.

5.2.2 Training, validation and testing set design

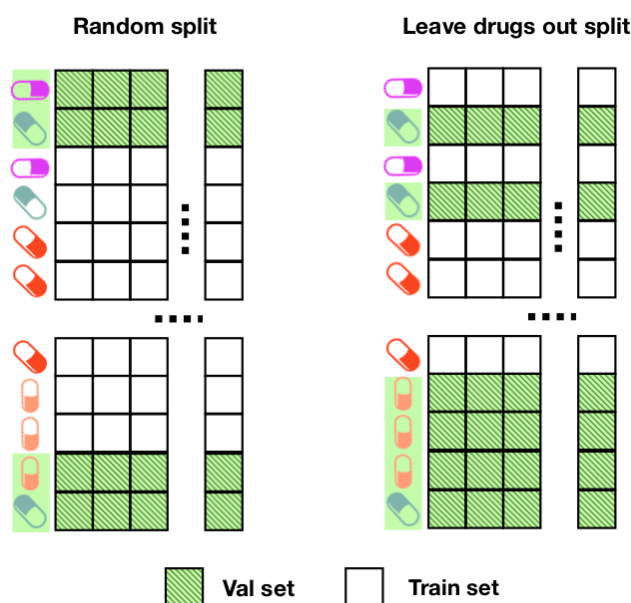


Figure 5-3 Considered cross-validation approaches. Random cross-validation results in validation and testing sets having profiles of drugs from training sets, which leads to inflated performance on validation and low generalisation abilities on unseen test data compared to leave-drug-out cross-validation.

Models were trained on the expression values of relevant genes, seven molecular descriptors values and 79 fingerprint values (340 features in total), calculated as detailed in Section 4.2.2. As specified in the section 'Training testing set design' of Chapter 4 (page 85), models were trained on matrix size of 340×1154 , and tested on 340×746 .

Figure 5-3 shows the two cross-validation strategies compared in this study. An optimal cross-validation strategy is crucial in order to accurately assess the performance of the model on unseen drugs, and therefore to evaluate its translational potential into real-world practice. For each cardiotoxicity label, its respective training and validation sets were preserved completely non-overlapping by drugs, with a leave-drug-out cross-validation strategy. To determine the generalisation ability of the proposed methods for novel drugs, I assessed the best-performing models on the selected testing set of 66 unique drugs. This strategy was compared with random cross-validation, frequently used in machine learning studies, including for drug safety.

5.2.3 Multi-label balancing

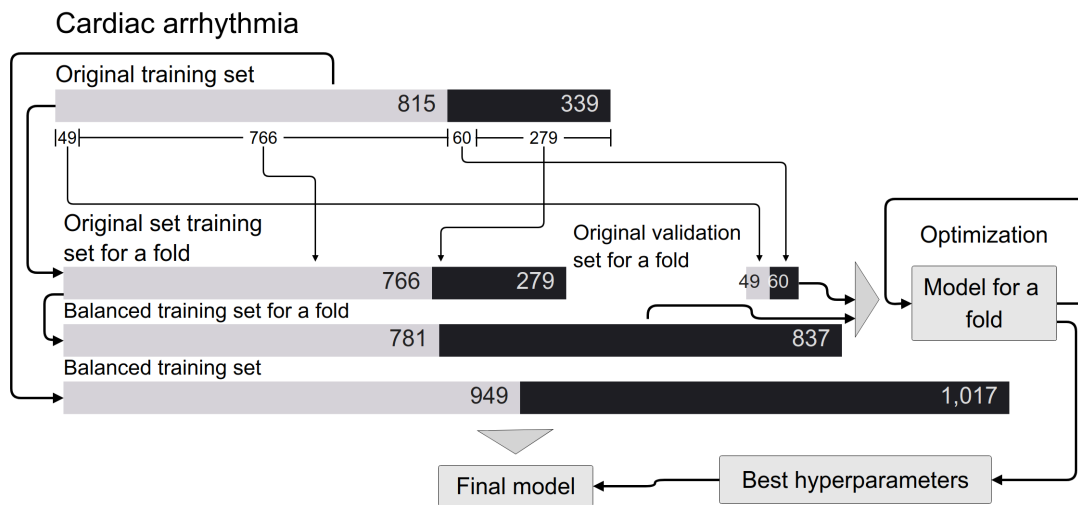


Figure 5-4 Multi-label balancing strategy illustrated with the fold 1 of 'cardiac arrhythmia'. For each cardiotoxicity form, an original training set was selected and subsequently split into validation and training sets for each fold. The training set for a fold was balanced, and used together with the original intact validation set, to optimise the model and find the best set of hyperparameters. Afterwards, the final model was trained on the balanced training set with the best set of selected hyperparameters.

I performed the synthetic minority over-sampling technique (SMOTE; [333]), implemented in the DMwR R package, to balance datasets. Figure 5-4 shows the balancing strategy for one cardiotoxicity type. The same strategy was applied for each fold and cardiotoxicity form. Balancing was only performed in the training set in cross-validation, in order to avoid over fitting at any stage.

5.2.4 Chain of classifiers with nested stacking

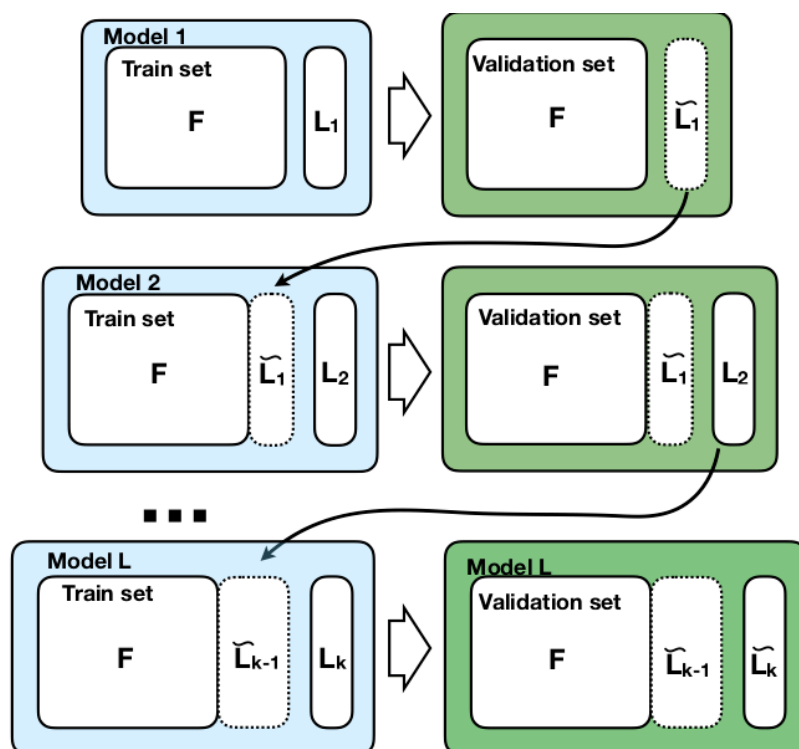


Figure 5-5 Used chain of classifiers with nested stacking model for multi-label prediction of drug cardiac safety. Each classifier L in the chain is trained on a set of F features and the $L-1$ labels predicted by the previous $L-1$ classifiers

Figure 5-5 outlines the employed model of chain of classifiers with nested stacking [334], which takes into account label dependencies. A chain of classifiers with nested stacking is a model that receives a feature vector and maps it to a set of labels. Each classifier in the chain is trained on a set of features together with the set of labels predicted by the previous classifiers. Nested stacking means that predictions for subsequent labels are made only for validation sets in each fold. This ensures that the classifiers do not predict values for samples that were included in the training for that model, therefore improving the generalisation abilities of such a classifier on unseen data.

I used cardiotoxicity types as labels, and gene expression values along with molecular descriptors and fingerprints as features. I used the following order of cardiotoxicity types obtained from MedDRA (see Chapter 5.2.1 Data collection page 82): 'Vascular disorders', 'Cardiac disorder signs and

symptoms', 'Cardiac arrhythmias', 'Heart failure', 'Coronary artery diseases', 'Pericardial disorders' and 'Myocardial disorders'. Labels were ordered by increasing number of drugs related to those side effects. Because the model for the first label will not receive information from any other cardiotoxicity types, I introduced 'Vascular Disorders' (which is also related to cardiac disorders) as the first disorder for prediction, in order to minimise the effect of the first position for 'Cardiac disorder sign and symptoms'. However, the accuracy of 'Vascular disorder' prediction was not used in the evaluation of the model performance and in selection of the best model.

To determine which algorithm accounts best for the observed data, I adapted several supervised classification algorithms widely used in bioinformatics:

- elastic net logistic regression (glmnet R library by Friedman et al. [335]). This model uses a logistic function to model the presence or absence of cardiac adverse reaction reports as a binary dependent label. The elastic net regularisation method linearly combines lasso [336] and ridge [337] penalty methods (which penalise the sum of absolute or squared values of coefficients, respectively) to select transcriptional, molecular descriptors and fingerprint features.
- random forest (ranger R library by Marvin et al. [338]). Random forest classifiers [339] combine a series of decision tree classifiers trained on different subsets of transcriptional, molecular descriptors and fingerprint features.
- gradient boosting logistic regression (gbm R library by Ridgeway et al. [340]). Gradient boosting methods [341] combine a series of logistic regression classifiers trained on different subsets of transcriptional, molecular descriptors and fingerprint features.
- categorical boosting (catboost R library by Prokhorenkova et al. [342]). Categorical boosting is a modification of the gradient boosting method, introduced in 2017. This method is also a combination of decision tree classifiers, but it allows handling categorical values and also uses new schema for selection of trees. As a result, it has been shown to outperform other gradient boosting models [342].

5.2.5 A set of individual classifiers

I used a set of individual classifiers as a binary relevance method. As in case of the chain of classifiers, each classifier is trained on a set of features, but in contrast to the chain of classifiers, it does not receive any previously predicted labels.

5.2.6 Optimisation schema

All models were optimised with Latin hypercube sampling of parameters (clhs R library by Roudier [343]) towards maximum Matthews correlation coefficient:

$$(5.1) \text{MCC} = \frac{\sum tp \times \sum tn - \sum fp \times \sum fn}{\sqrt{(\sum tp + \sum fp)(\sum tp + \sum fn)(\sum tn + \sum fp)(\sum tn + \sum fn)}},$$

where tp (true positives) and tn (true negatives) are the number of compounds with cardiotoxicity reports and the number of compounds without cardiotoxicity reports predicted correctly, respectively, and fp (false positives) and fn (false negatives) are the number of compounds without cardiotoxicity reports and the number of compounds with cardiotoxicity reports predicted incorrectly, respectively. An $\text{MCC}=0$ indicates that the prediction is not better than a random prediction, an $\text{MCC}=1$ indicates perfect prediction or total agreement, and an $\text{MCC}=-1$ indicates total disagreement.

I first trained models with five-fold cross-validation selected to leave drugs out to compensate for overfitting, and to receive more robust performance metrics. Once trained, to predict the cardiac safety of any unseen chemicals, the model only receives as inputs their transcriptional features, molecular descriptors, and fingerprints.

5.2.7 Model evaluation

In addition to MCC, the following metrics were used to evaluate model performance:

$$(5.2) \text{accuracy} = \frac{\sum tp + \sum tn}{\sum tp + \sum tn + \sum fp + \sum fn},$$

$$(5.3) F1 = 2 \times \frac{precision \times recall}{precision + recall},$$

where $precision = \frac{\sum tp}{\sum tp + \sum fp}$ and $recall = \frac{\sum tp}{\sum tp + \sum fn}$, with tp (true positives), tn (true negatives), fp (false positives) and fn (false negatives) defined as before. Accuracy represents the ratio of correctly predicted drugs to a total number of drugs. Precision equals the fraction of correctly predicted unsafe compounds in all compounds predicted as unsafe, whereas recall shows the sensitivity of a model and equals the fraction of correctly predicted unsafe compounds out of all real unsafe compounds. An F1 score of 1 implies perfect precision and recall, while $F1=0$ denotes the worst accuracy performance.

Model performance was also evaluated against the area under the receiver operating characteristic (ROC) curve (AUC):

$$(5.4) AUC = \int_{-\infty}^{\infty} TPR(T)(-FPR'(T))dT,$$

where TPR is the true positive rate (identical to recall) and FPR is the false positive rate. AUC measures the diagnostic ability of a predictor, where an $AUC=0.5$ indicates that the prediction is not better than a random prediction, and an $AUC=1$ indicates perfect prediction. The pROC R library by Robin et al. [344] was used to calculate AUC values for the classifiers.

Table 5-1 Performance of multi-label classification models, trained on transcriptional profiles and molecular descriptors and fingerprints of drugs, on validation and testing sets. Reported metrics are area under the receiver operating characteristic curve (AUC; upper value) and Matthew correlation (MCC; lower value). Validation and testing set performances are shown in the upper and the lower cells of each row, respectively. The best performance is shown in bold. RF: random forest; ELNET: elastic net logistic regression; GBM: gradient boosting machines; CATBOOST: categorical boosting; Cardiac dis S&S: MedDRA term for cardiac disorders signs and symptoms; dis: disorders.

Cardiotoxicity form	Set		Cardiac dis S&S	Cardiac arrhythmias	Heart failures	Coronary artery dis	Pericardial dis	Myocardial dis	Mean $\pm$ 0.5 SD
Leave drug out cross-validation strategy									
Chain of classifiers with nested stacking									
Dual feature set									
RF	Validation	AUC	63	88	79	91	64	88	79 $\pm$ 6
		MCC	0.21	0.50	0.29	0.68	0.13	0.46	0.38 $\pm$ 0.10
	Testing	AUC	70	64	58	54	83	87	66$\pm$9
		MCC	0.33	0.37	0.04	0.10	0.14	-0.08	0.15$\pm$0.9
ELNET	Validation	AUC	62	67	65	55	71	65	64 $\pm$ 3
		MCC	0.20	0.23	0.19	-0.03	0.6	0.37	0.17 $\pm$ 0.07
	Testing	AUC	67	65	63	49	66	70	63 $\pm$ 4
		MCC	0.28	0.28	-0.08	-0.08	-0.03	0.01	0.06 $\pm$ 0.09
GBM	Validation	AUC	67	87	74	90	60	86	77 $\pm$ 6
		MCC	0.29	0.55	0.35	0.66	0.16	0.54	0.42 $\pm$ 0.09
	Testing	AUC	66	62	51	55	60	64	60 $\pm$ 3
		MCC	0.22	0.26	-0.08	0.03	0.12	-0.04	0.08 $\pm$ 0.07
CATBOOST	Validation	AUC	64	84	64	88	65	53	70 $\pm$ 7
		MCC	0.30	0.68	0.32	0.65	0.24	0.02	0.37 $\pm$ 0.13
	Testing	AUC	67	63	59	57	67	49	60 $\pm$ 3
		MCC	0.27	0.18	-0.07	0.09	0.35	0.0	0.14 $\pm$ 0.08

Transcriptional features only									
RF	Validation	AUC	63	83	82	90	61	89	78±6
		MCC	0.24	0.47	0.4	0.66	-0.02	0.62	0.4±0.13
	Testing	AUC	69	65	58	64	71	67	66±2
		MCC	0.22	0.19	0	0.21	0.11	0.11	0.14±0.04
Descriptors and molecular fingerprints only									
RF	Validation	AUC	66	83	82	86	48	81	74±7
		MCC	0.24	0.55	0.37	0.51	-0.08	0.19	0.3±0.12
	Testing	AUC	63	72	42	56	62	58	59±5
		MCC	0.21	0.22	0.01	0.16	-0.08	-0.07	0.08±0.07
Set of independent binary classifiers									
RF	Validation	AUC	57	62	61	61	57	53	58±2
		MCC	0.14	0.12	0.06	-0.03	0.0	0.17	0.08±0.04
	Testing	AUC	64	63	59	59	45	58	58±3
		MCC	0.21	0.18	-0.02	0.12	-0.01	-0.01	0.08±0.05
ELNET	Validation	AUC	60	58	61	55	72	64	62±3
		MCC	0.12	0.15	0.10	-0.08	0.06	0.17	0.09±0.05
	Testing	AUC	65	67	61	52	64	65	62±3
		MCC	0.19	0.25	-0.03	-0.08	-0.04	-0.05	0.04±0.07
GBM	Validation	AUC	59	61	50	49	58	66	57±3
		MCC	0.11	0.22	-0.02	-0.03	-0.14	0.37	0.08±0.09
	Testing	AUC	67	61	54	57	50	63	59±3
		MCC	0.21	0.20	-0.02	0.04	-0.03	0.05	0.08±0.05
CATBOOST	Validation	AUC	62	60	61	56	54	70	60±3
		MCC	0.23	0.26	0.20	0.28	0.07	0.37	0.24±0.05
	Testing	AUC	72	66	66	60	55	61	63±3
		MCC	0.37	0.22	0.0	0.11	0.0	0.06	0.13±0.07

Random cross-validation strategy									
Chain of classifiers with nested stacking									
RF	Validation	AUC	92	96	93	95	94	92	94±1
		MCC	0.70	0.77	0.53	0.72	0.44	0.26	0.57±0.10
	Testing	AUC	68	62	52	50	73	79	64±6
		MCC	0.21	0.28	-0.09	-0.04	0.27	-0.07	0.09±0.09
Set of independent binary classifiers									
RF	Validation	AUC	90	88	74	83	86	77	83±3
		MCC	0.67	0.61	0.25	0.57	0.56	0.16	0.47±0.11
	Testing	AUC	66	63	59	56	45	58	58±4
		MCC	0.20	0.20	0.03	-0.09	-0.01	-0.01	0.05±0.06

5.3 Results

Table 5-1 summarises the performance of the different algorithms explored as a chain of classifiers (where each classifier passes predicted values to the next classifier in a chain), and as a set of individual classifiers, under both leave-drug-out cross-validation and for random cross-validation. A detailed description of these results is presented in the sections below.

5.3.1 Chain of classifiers versus a set of individual predictors

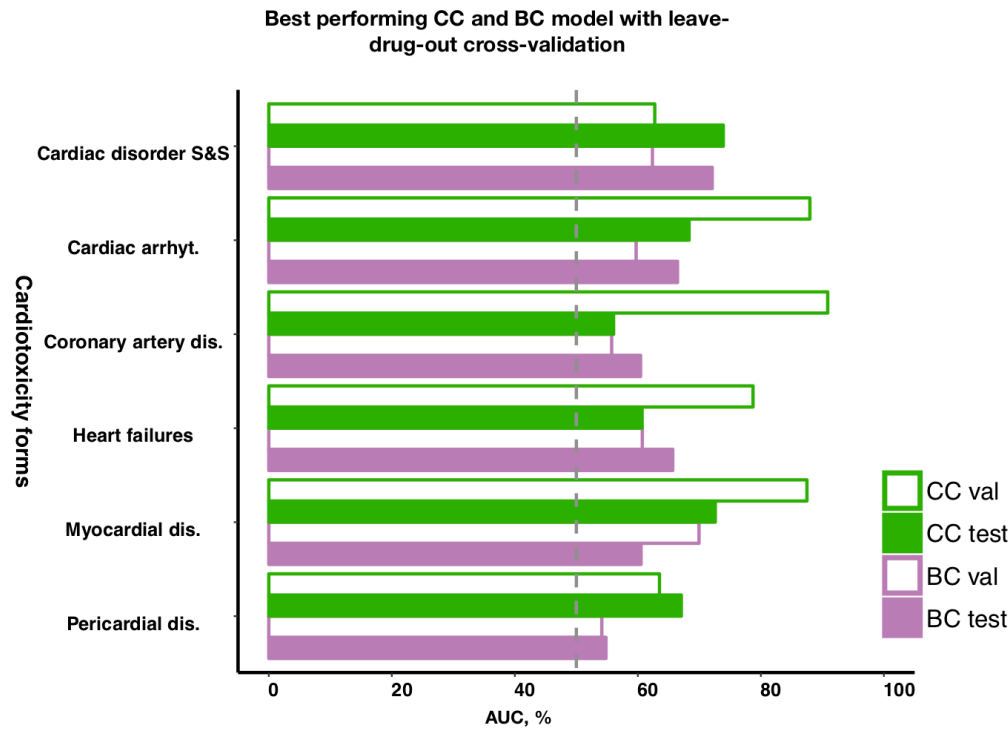


Figure 5-6 Prediction of multiple forms of cardiotoxicity using the chain of classifiers with nested stacking (CC) versus a set of binary classifiers (BC). (A) AUC of best performing CC and BC in safety drug prediction on validation (val) and testing (test) sets under leave-drug-out cross-validation, for each independent cardiotoxicity label.

Figure 5-6 shows the comparison of the best chain of classifier (random forest) and the best set of individual classifiers (categorical boosting) under leave-drug-out cross-validation. In general, the chain of classifiers outperforms the best set of individual classifiers, with the best results obtained for the chain of random forest classifiers, with an average AUC of 79% and an average MCC of 0.38 across all cardiotoxicity forms on validation, and 66% and 0.15 on testing (Table 5-1 and Figure 5-6). The second best result was

obtained with a chain of gradient boosting classifiers, with average AUC of 77% and average MCC of 0.42 on validation, and 60% and 0.08 on test set. The chain of categorical boosting classifiers showed average AUC of 70% and average MCC of 0.37 on validation, with average AUC of 60% and average MCC of 0.14 for testing. Finally, elastic net demonstrated the most modest performance among the trained set of chain classifiers, achieving an average AUC of 64% and average MCC of 0.17 on validation, with average AUC of 63% and average MCC of 0.06 for testing.

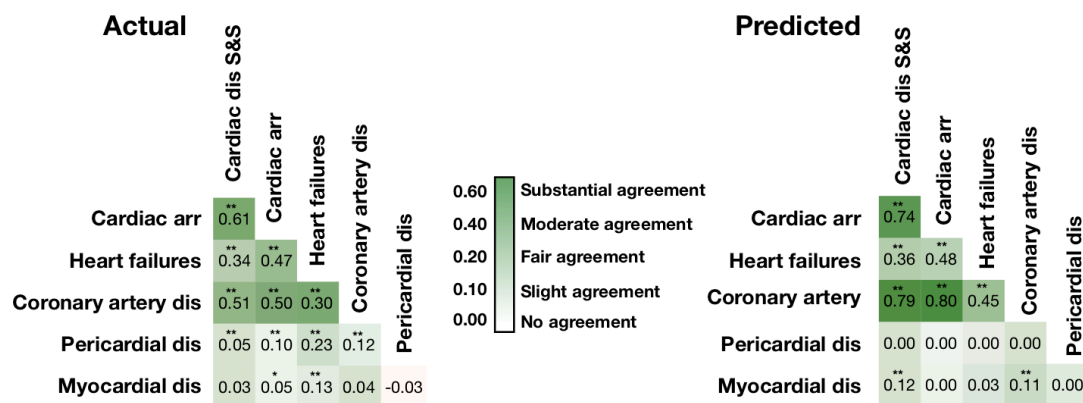


Figure 5-7 Levels of association measured between cardiotoxicity forms in the testing set in terms of Cohen's Kappa (left). Levels of association between predicted cardiotoxicity forms are measured in terms of Cohen's Kappa (right). Symbols represent statistical significance level by Z-test (* $p < 0.05$, ** $p < 0.01$). 'Cardiac dis S&S': MedDRA term for cardiac disorder signs and symptoms; 'dis': disorders.

Figure 5-7 shows the level of association for predicted cardiotoxicity forms with the best performing chain of classifiers model, which demonstrate similarity to the ones shown in the original data. This was particularly clear for the predominant associations between cardiac disease signs and symptoms, cardiac arrhythmias and coronary artery disease.

At the same time, for a set of individual classifiers, the best set of categorical boosting classifiers only obtained an average AUC of 60% and an average MCC of 0.24 across all cardiotoxicity types on validation, with an average AUC of 63% and average MCC of 0.13 on testing (Table 5-1 and Figure 5-6). The second best was logistic regression with elastic net regularisation (AUC of 62% and MCC of 0.09 on validation and AUC of 62% and MCC of 0.04 on

testing). Interestingly, a set of independent gradient boosting classifiers was almost as accurate as a chain, with AUC of 59% and MCC of 0.08 vs AUC of 60% and MCC of 0.07. A set of individual random forest classifiers only achieved 58% AUC and 0.08 MCC on testing.

Therefore, while cardiotoxicity types were predicted with different accuracies, the inclusion of information between cardiotoxicities improved prediction accuracy for all cardiotoxicity types.

5.3.2 Random vs leave-drug-out cross-validation

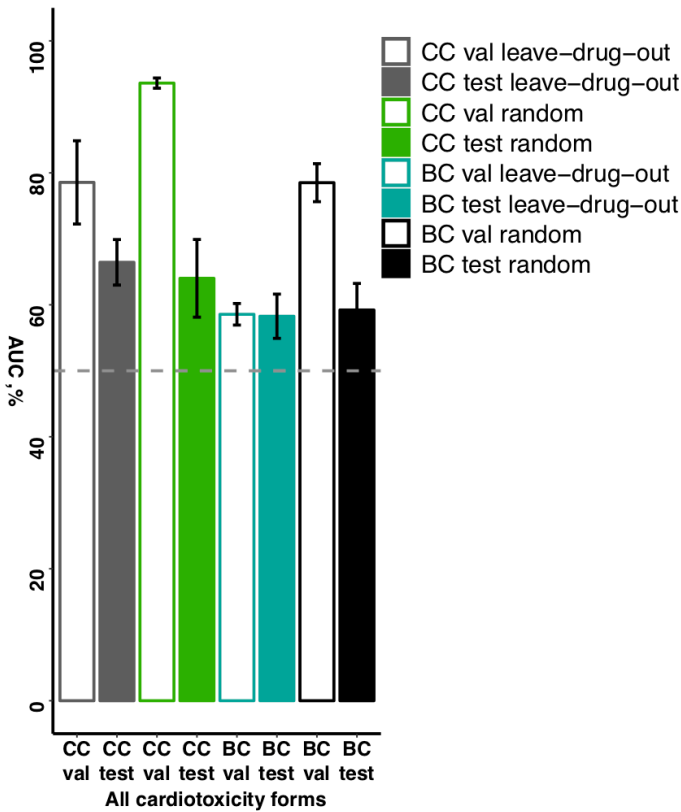


Figure 5-8 Comparison between leave-drug-out vs random cross-validation strategies. Average performance across all labels is shown for the best performing chain of classifiers (CC) model, trained with leave-drug-out vs random cross-validation strategies, and for a set of independent binary classifiers (BC). Random cross-validation significantly inflates the accuracy of the trained models compared to leave-drug-out validation.

Figure 5-8 shows the AUC values for the chain of random forest classifiers (CC) and a set of individual random forest classifiers (BC), both trained with either leave-drug-out or random cross-validation strategies. When validated and optimised with random cross-validation, both models demonstrate almost perfect accuracy on validation (averages for all cardiotoxicity types: AUC=94%, MCC=0.57 for chain of classifiers; AUC=83%, MCC=0.47 for independent predictors; Figure 5-8 and Table 5-1). Although apparently outperforming the models trained with leave-drug-out cross-validation, models

trained with random cross-validation were however less accurate when predicting cardiotoxicity types of previously unseen drugs (test datasets in Figure 5-8 and Table 5-1). Quantitatively, the chain of classifiers trained and optimised with a random cross-validation strategy exhibited an AUC of 64% and MCC of 0.09, compared to an AUC of 66% and MCC of 0.15 for leave-drug-out cross-validation (average values for all cardiotoxicity types). Similarly, a set of individual classifiers with random cross-validation showed an AUC of 58% and MCC of 0.09, compared to an AUC of 58% and MCC of 0.05 in the case of leave-drug-out cross-validation.

5.3.3 Dual vs individual feature sets

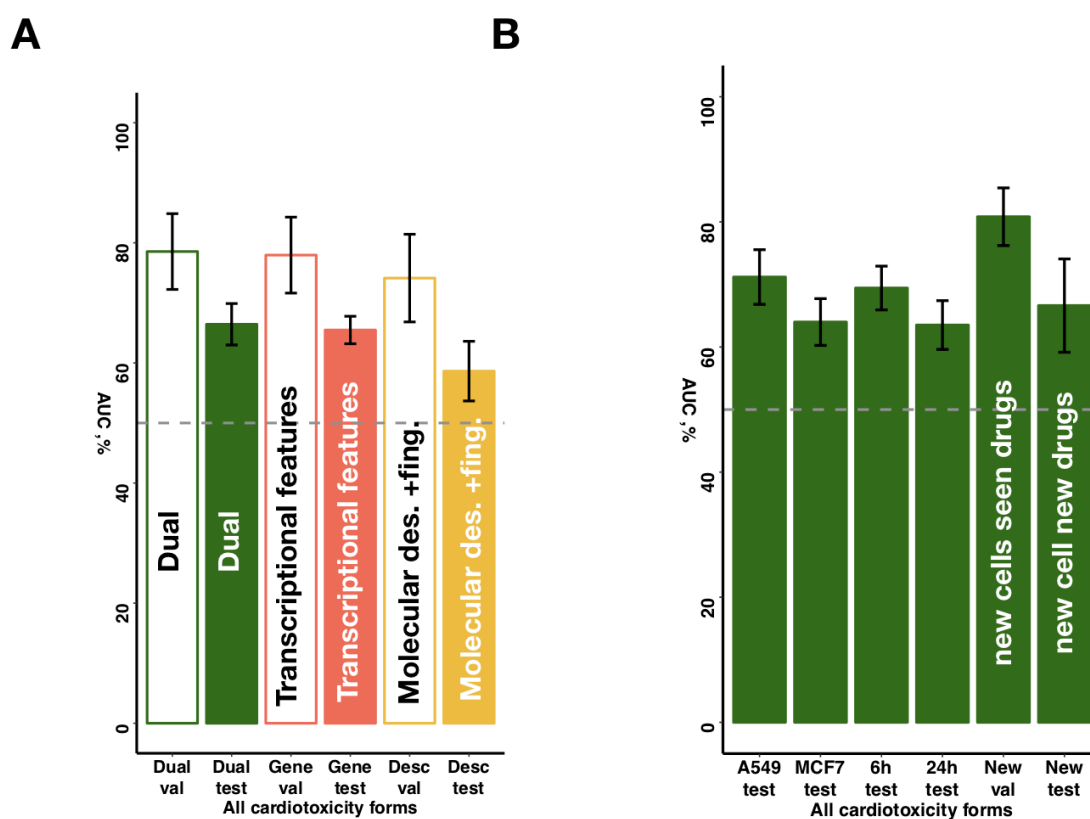


Figure 5-9 Sensitivity of predictive accuracy to feature types, cell lines and incubation times. **(A)** Merging of molecular descriptors and fingerprints and transcriptional features results in increased average performance across all drug classes. **(B)** The best performing model demonstrates similar predictive accuracy across different cell lines (A549, MCF7) and incubation times (6 and 24 hours). For new cell lines (PHH, SKB, SKM1, A673), the model discriminates more accurately seen drugs with unseen cell lines ('new cells seen drugs') than unseen drugs ('new cells new drugs').

To further explore the predictive value of individual feature sets, I compared for the entire dataset the proposed chain of random forest classifiers with nested stacking separately trained on each feature set (transcriptional versus molecular descriptors).

Figure 5-9A summarises the performance of models trained only on one feature type and on the dual set. The dual set shows on validation an AUC of 79% and MCC of 0.38 versus an AUC of 74% and MCC of 0.30 for molecular descriptors and fingerprints, or an AUC of 78% and MCC of 0.40 for transcriptional features only (Figure 5-9, Table 5-1). Similarly, the dual transcriptomic and molecular classifier is more accurate on testing. The model trained only on molecular descriptors and fingerprints achieves 66% AUC and 0.14 MCC, being 59% AUC and 0.08 MCC when trained only for transcriptional features.

Following the observation of the clustering of samples according to cell types and incubation times (Figure 4-6), I decided to explore the sensitivity of the best performing model to those sample characteristics. Figure 5-9B illustrates that the best model tends to classify slightly more accurately drugs profiled with A549 compared to MCF7 cell lines, and drugs incubated for 6 compared to 24 hours (Supplementary Table 5-1). To further investigate robustness to unseen cell lines, I used the best performing model to predict cardiac safety for transcriptional profiles on four additional cell lines: PHH, SKB, SKM1, A673 (Supplementary Table 5-1). For already seen drugs, the model discriminates with high accuracy (average AUC=81%, MCC=0.32), indicating that it is learning the information on cardiotoxicity regardless of the cell line used. For unseen drugs profiled with totally new cell lines, the model exhibits a moderate decrease in performance (average AUC=0.67 and MCC=0.04 vs 0.66 and MCC=0.14), suggesting good generalisation properties of the proposed model.

5.3.4 Predictive accuracy across drug classes and cardiotoxicity forms



Figure 5-10 Predictive accuracy across drug classes, cardiotoxicity forms and feature types. The best performing model trained on either molecular descriptors and fingerprints, transcriptional features or both demonstrates different performance in predicting different types of agents. For each drug class, the top bar plot shows the AUC of the best predictor,

and the respective bottom bar plot the aggregated number of safe and unsafe drugs in the training and test sets. CV: cardiovascular agents; CNS: central nervous system agents.

Figure 5-10 shows the performance of the best model (chain of random forest classifiers with nested stacking) across cardiotoxicity forms and drug classes (Supplementary Table 5-1). For example, for antineoplastic drugs, the best model predicted 'pericardial disorders' more accurately (Supplementary Table 5-1; AUC=75% and MCC=0.47) than the average across all drugs, also achieving high accuracy in the prediction of 'heart failure' (AUC=76% and MCC=0.13). 'Cardiac disorder signs and symptoms' were also more accurately predicted in the case of anti-inflammatory drugs compared to other cardiotoxicity types (Supplementary Table 5-1; AUC of 86% and MCC of 0.43).

On the contrary, predictions for 'pericardial disorders' were close to random guessing for central nervous system agents (AUC of 53% and MCC of -0.03) and for 'cardiac disorders signs and symptoms' for cardiovascular agents (AUC of 52% and MCC of -0.16). Some of these cases of limited performance may be partially explained by a small frequency of side effects for specific drug types. For example, only 16 out of 56 total cardiovascular agents in the collected dataset are known to cause cardiac arrhythmias and coronary artery diseases (Figure 5-10). However, for antineoplastic drugs the frequency of coronary artery disease (7/33) is the same than for cardiac arrhythmias (7/33) and bigger than for heart failure (3/33), but the predictor performs better in predicting the latter two than coronary artery disease, indicating a possible dependency on feature types.

Interestingly, the model trained individually on molecular descriptors and fingerprints demonstrated higher accuracy than the transcriptomic or dual predictors in safety predictions of 'cardiac arrhythmias' for cardiovascular, antineoplastic and central nervous system agents, but not for anti-inflammatory agents (Figure 5-10). For central nervous system agents, in general, the molecular descriptor-based predictor is more accurate. Notably,

for 'coronary artery disorders', a combination of the two feature types leads to a decrease in accuracy compared to individual feature set predictors.

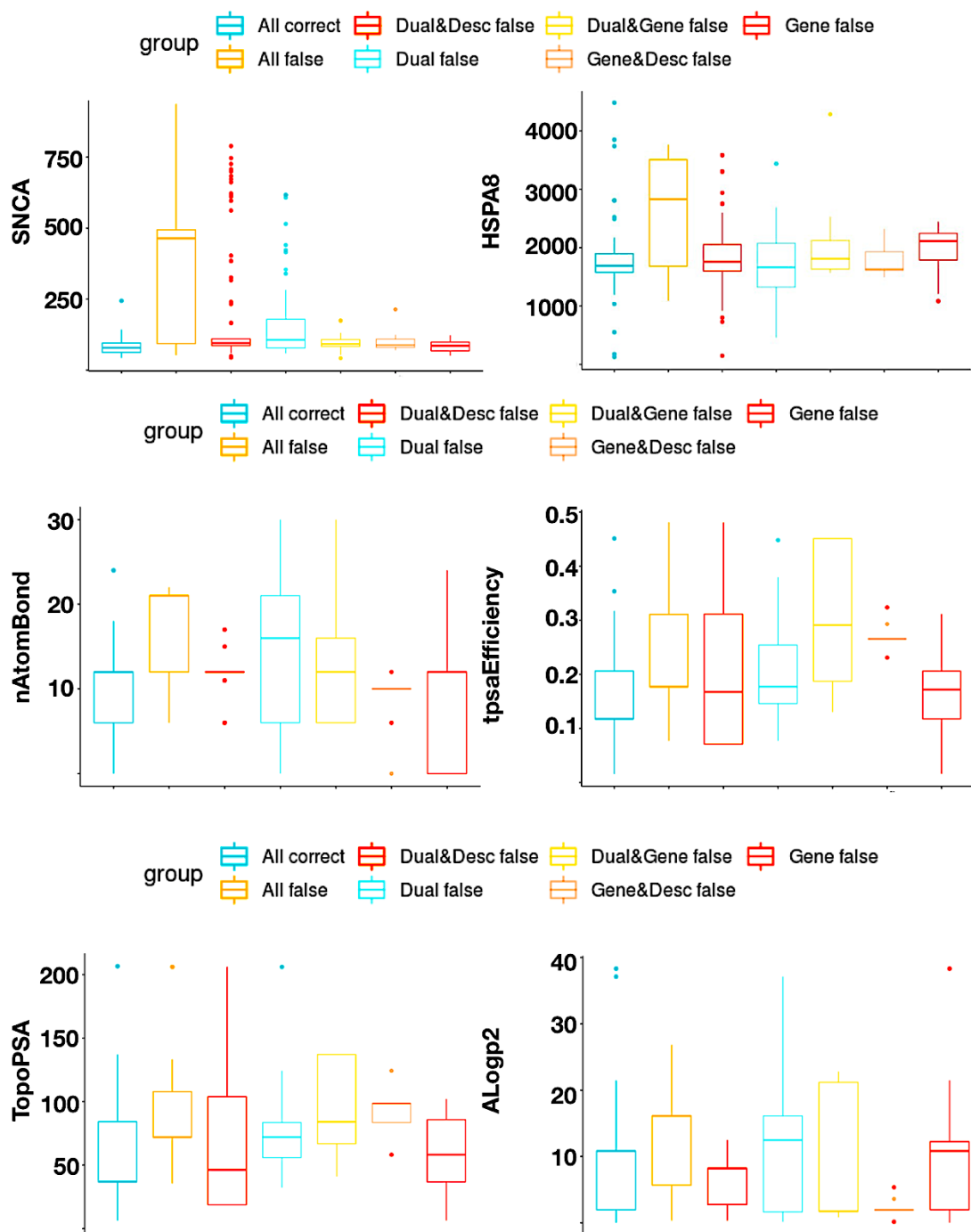


Figure 5-11 Values of SNCA, HSPA8, nAtomBond, tpsaEfficiency, TopoPSA and ALogp2 features, for drugs assigned to seven prediction groups according to prediction accuracy. 'All correct' and 'all false' indicate samples predicted correctly and incorrectly by all three predictors, respectively. Equivalently, the rest of groups represent samples incorrectly predicted by the dual predictor ('Dual'), by the predictor trained only on transcriptional features ('Gene') and by the predictor trained on molecular descriptors only ('Desc').

Figure 5-11 shows an example of transcriptional features and molecular descriptors associated with prediction accuracy. For instance, drug samples

with higher expression values of Alpha-Synuclein (SNCA) or Heat Shock Protein 8 (HSPA8) genes were more often predicted incorrectly by all three predictors. Drugs with a higher number of atomic bonds (nAtomBond) were classified less correctly when trained on the combined features set or only on molecular descriptors. The same was observed for drugs with higher values of polar surface area to molecular size ratio (tpsaEfficiency) or topological polar surface area based on fragment contributions (TopoPSA).

5.4 Discussion

The proposed method for multi-label prediction of drug-induced cardiotoxicity (chain of random forest classifiers with nested stacking) achieves a 79% average AUC across the six main types of cardiotoxicity on a leave-drug-out cross-validation strategy on 291 drugs (8,237 samples). Further validation of the method on new and completely unseen 66 drugs (1,696 samples), with multiple mechanisms of action, demonstrates that the proposed model holds high generalisation abilities compared to sets of individual classifiers. Models trained with leave-drug-out cross-validation were able to discriminate between safe and unsafe drugs with a 66% average AUC and MCC of 0.15 across all cardiotoxicity types. The model demonstrates higher accuracy for specific adverse drug effects and type of agents, and in particular for pericardial disease, cardiac disease and symptoms, heart failure and myocardial disease for antineoplastic, anti-inflammatory, cardiovascular and central nervous system agents, respectively. These results suggest the translational potential of the proposed approach towards applications in a pre-clinical context.

The proposed method takes advantage of the chain of classifiers approach. This approach significantly outperformed binary classification approaches that treat each label independently, with an improvement of 5% in terms of AUC (from AUC 63% to 66%) and 15,38% MCC (0.13 to 0.15). This highlights the importance of incorporating information about related adverse reactions in predicting drug safety. While demonstrating good generalisation abilities on unseen data, the model showed different performance across cardiotoxicity

types depending on the type of agents. Cardiac arrhythmia-related safety was predicted more accurately for cancer and anti-inflammatory agents than for cardiovascular and central nervous system agents. This might be improved by the introduction of information more relevant to specific mechanisms of arrhythmogenesis.

The accumulated body of evidence suggests that gene expression signatures alone could also be used as a biomarker of cell response to drugs [345,346]. Those results, in line with previous studies [291], demonstrate that coupling of transcriptional profiles and molecular descriptors improves the predictive power of algorithms. Thus, the combination of both feature types indeed increases the mean accuracy of a chain of classifiers.

Previous research [291,292] reported good accuracy in the prediction of multiple adverse reactions, including liver disorders, gastrointestinal disorders, myocardial infarction and cardiomyopathy. However, such approaches were tested using drugs included in the training, rather than in an independent dataset as performed in this study. They also neglected existing dependencies between various forms of cardiotoxicity, as I have demonstrated in this chapter, where I also show that retaining information about cardiotoxicity dependencies results in greater prediction accuracy. At the same time, I show that models trained with random cross-validation may display a significantly inflated performance on validation, nevertheless becoming less accurate when predicting previously unseen drugs. In general, the leave-drug-out cross-validation strategy demonstrated a more robust performance compared to random cross-validation, the latter evidencing inflated accuracy metrics, which in turn may complicate model optimisation and overstate their expected generalisation ability. The proposed chain of classifiers model with nested stacking has indeed better generalisation for multi-label predictions than previous models, such as that by Wang and colleagues [291], and demonstrates a superior performance compared to models based on sets of individual classifiers.

The proposed model can predict both acute (cardiac arrhythmias) and chronic (coronary artery disorders and heart failures) effects of drugs, based on clinical human responses collected via SIDER. Chronic drug-induced cardiac changes are often irreversible and their effects are delayed. Therefore, their prediction poses a great challenge, as they would require long-term animal experimentation, together with translational challenges from animal to human [347]. Another key advantage of the approach is the use of perturbation databases such as LINCS and Connectivity map. They constitute great resources for preclinical applications [348], and they have even been used extensively to identify novel drug candidates that confirmed their effectiveness experimentally [349]. Providing a cheaper alternative to animal models, computational methods that integrate transcriptional responses of drugs and their molecular characteristics, such as the one I propose here, could be used prior to animal experiments to identify drug cardiotoxicity early in the pre-clinical phase of drug discovery.

Using transcriptional signatures, molecular descriptors and fingerprints, the general methodology proposed in this study could be further applied to other tissue-specific side effects and organs. I presume that this approach can be also extended to other areas including drug target prediction, where information about the multi-label properties of drugs or multi-target properties plays a vital role [350].

5.5 Conclusions

In this chapter, I have presented a machine-learning approach for simultaneous prediction of six forms of cardiotoxicity. The achieved overall good performance suggests that gene expression changes can be used as biomarkers of drug-induced cardiotoxic response. However, while the proposed model shows relatively good accuracy in discriminating between safe and drugs that can cause cardiac adverse reactions, many aspects of the underlying mechanisms of cardiotoxicity are yet to be well understood. In the next chapters, I will present a mechanistic model that incorporates information

about expression changes to investigate the effects of hERG trafficking disruption by drugs on cardiac electrophysiology.

6 Computational modelling and simulation of hERG protein dynamics in human cardiomyocytes

6.1 Introduction

At the end of Chapter 5, I described how machine learning can be used to predict the cardiotoxicity of drugs using transcriptional and structural information. While being reasonably accurate in the prediction of outcomes, machine learning approaches have a limited ability to identify the underlying mechanisms of a drug's off-target actions. An alternative for the latter is the use of mechanistic-based models, which, for example, have demonstrated high accuracy in the prediction and explanation of clinical pro-arrhythmia due to acute ion channel inhibition [290,311,351]. However, the extension of mechanistic-based models to other types of more complex drug off-target interactions, such as drug-induced trafficking disruption, is not straightforward. It requires the development of suitable models to capture these interactions in order to augment the predictive value of *in silico* approaches.

In this chapter, I describe the development of a new mechanistic model of hERG channel trafficking and its application in investigating the effect of abnormalities on trafficking on cellular cardiac electrophysiology. I start with a brief summary of the key stages of hERG trafficking, which were discussed in detail in Chapter 2, in light of cell-drug interactions ("hERG trafficking inhibition" section, Page 23). Afterwards, I formulate the model and describe its calibration based on experimental data. Finally, I present the effect of changes in hERG trafficking on action potential models of human iPSC derived cardiomyocytes.

6.2 Materials and Methods

6.2.1 Characterisation of hERG trafficking stages and key regulatory proteins

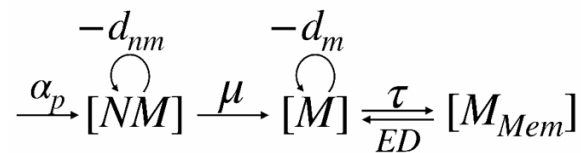
The hERG channel is encoded by the KCNH2 gene, also known as hERG, located on chromosome 7. The biogenesis, trafficking and degradation of hERG can be subdivided into the following steps (Figure 2-4):

- 1) Synthesis of non-mature hERG, which includes the synthesis of pre-mRNA in the nucleus (transcription), followed by splicing, capping, cleavage and polyadenylation, and its transport to the endoplasmic reticulum, where it will be translated to the polypeptide.
- 2) A fraction of non-mature hERG protein undergoes maturation, and the remainder are degraded.
- 3) Maturation of hERG, which includes the transport of the non-mature hERG (polypeptide) in the COPII-mediated vesicle to the Golgi apparatus with Hsp70 and Hsp90 chaperones attached, where the protein will be glycosylated.
- 4) A fraction of the mature hERG protein is transported to the membrane, and the rest is degraded.
- 5) Exocytosis of mature hERG (glycosylated form) to membrane.
- 6) Endocytosis of mature hERG, where it can be recycled or degraded.

It should be noted that some of the processes listed above have been shown to be integrated with each other and occur almost simultaneously, such as the pre-processing and transcription of mRNA [352].

6.2.2 Formulation of the hERG trafficking model

The mathematical formulation of hERG dynamics described above can be represented as the following model:



where variables $[NM]$, $[M]$, $[M_{Mem}]$ respectively represent protein densities of non-mature, mature and membrane hERG forms. Parameter α_p is the non-mature protein synthesis rate, μ is the maturation rate or transformation of non-mature to mature form, τ is the rate of transport to the cell membrane, ED

is the rate of endocytosis or stability on the membrane, and d_{nm} and d_m are the non-mature and mature protein degradation rates, respectively (all parameters positive). The model above can be described as a set of ordinary differential equations (ODEs). The equations read:

$$(6.1) \frac{d[NM]}{dt} = \alpha_p - \mu \cdot [NM] - d_{nm} \cdot [NM],$$

$$(6.2) \frac{d[M]}{dt} = \mu \cdot [NM] - \tau \cdot [M] + ED \cdot [M_{Mem}] - d_m[M],$$

$$(6.3) \frac{d[M_{Mem}]}{dt} = \tau \cdot [M] - ED \cdot [M_{Mem}],$$

with equations (6.1)-(6.3) respectively describing the time evolution of hERG non-mature, mature and mature on the cellular membrane forms.

6.2.3 Implementation of the model

For parameter optimisation, I implemented the model in Matlab (R2017a version; [353]) and solved it with the 'ode15s' function [354]. The rest of simulations were performed in R (3.4 version [355]) and solved with the 'ode' function with the default integrator 'Lsoda' from the deSolve package [356]. The ggplot2 R library was used to visualise the results [357].

6.2.4 Calibration of the model

6.2.4.1 Characterisation of *in vitro* hERG trafficking experiments

Table 6-1 Experimental non-mature ($[NM]$) and total mature ($[M] + [M_{Mem}]$) hERG protein levels with maturation (μ), trafficking (τ) and endocytosis (ED) active, but synthesis blocked ($\alpha_p=0$). Protein levels were measured using radioactive labelling in HEK293 cells at 37°C. Data adapted from [358]. SD: standard deviation; n: number of independent experiments; a.u.: arbitrary units.

Times, h	$[NM]$ (mean $\pm$ SD, n=3), a.u.	$[M]+[M_{Mem}]$ (mean $\pm$ SD, n=3), a.u.
0	1.00 $\pm$ 0.00	0.00 $\pm$ 0.00
2	0.47 $\pm$ 0.02	0.32 $\pm$ 0.30
4	0.31 $\pm$ 0.18	0.43 $\pm$ 0.06
8	0.15 $\pm$ 0.06	0.26 $\pm$ 0.08

22	0.00±0.02	0.01±0.00
----	-----------	-----------

Ficker and colleagues [358] performed the analysis and quantification of the rate of non-mature hERG protein decay and mature protein synthesis for 24 hours using pulse-chase protocols (Table 6-1). They performed monitoring of hERG maturation in HEK293 cells incubated at 37°C using radioactive labelling. Cells were incubated in a [³⁵S]-methionine/cysteine medium for 30 minutes. This radioactive label is incorporated into newly synthesised hERG protein during incubation, and can be detected using phosphorimaging. Cells were lysated after 0, 2, 4, 8 and 22 hours to measure changes in the labelled proteins (disappearance of the non-mature hERG and appearance and disappearance of mature hERG).

Table 6-2 Experimental total mature ($[M] + [M_{Mem}]$) hERG protein with trafficking (τ) and endocytosis (ED) active, but maturation and protein synthesis blocked ($\alpha_p=0$ and $\mu = 0$). Protein levels were measured in HEK293 cells incubated at 37°C with Western blotting using immunodetection. Data adapted from [359]. n: number of independent experiments. a.u.: arbitrary units. Standard deviation values not available.

Times, h	[M]+[M _{Mem}] (mean, n=3), a.u.
0	1.00
2	0.82
4	0.72
6	0.54
8	0.57
11	0.30
24	0.08

Table 6-2 shows the data for a second experiment by Ke and colleagues [359]. For the quantification of degradation of the mature wild type and three trafficking hERG mutant channels, the authors used pulse-chase protocols and monitored HEK293 cells for 24 hours. Firstly, the cells were incubated at 27°C, then returned to 37°C and exposed to Brefeldin A, an agent that prevents the transport of proteins from the endoplasmic reticulum to the Golgi

apparatus, where they mature. Cells were lysated after 0, 2, 4, 6, 8, 11 and 24 hours to measure the expression of the fully glycosylated (mature) hERG by hybridisation with an anti-haemagglutinin antibody, followed by hybridisation with an antibody conjugated to a fluorescent dye and detection by an infrared imaging system. Protein levels were normalised to anti- β -actin antibody.

Table 6-3 Experimental non-mature ($[NM]$) and total mature ($[M] + [M_{Mem}]$) hERG protein levels with maturation (μ), trafficking (τ) and endocytosis (ED) active, but synthesis blocked ($\alpha_p=0$). Protein levels were measured in HEK293 cells incubated at 37°C using a radioactive detection system. Data adapted from [360]. SD: standard deviation; n: number of independent experiments; a.u.: arbitrary units.

Times, h	$[NM]$ (mean $\pm$ SD, n=5), a.u.	$[M]+[M_{Mem}]$ (mean $\pm$ SD, n=5), a.u.
0	1.00 $\pm$ 0.00	0.00 $\pm$ 0.00
1	0.80 $\pm$ 0.34	0.15 $\pm$ 0.06
2	0.56 $\pm$ 0.00	0.29 $\pm$ 0.02
4	0.51 $\pm$ 0.16	0.47 $\pm$ 0.14

Table 6-3 provides quantified hERG levels in a third experiment [360]. The authors used a pulse-chase protocol to measure non-mature and total mature hERG levels in HEK293 cells incubated at 37°C. The cells were transfected with SNAP-hERG plasmids, which react with a fluorescent label, thus allowing their detection. To initiate the experiment, non-fluorescent labels were added to block any already expressed hERG proteins with SNAP-tags, and further protein expression was blocked with cycloheximide. The cells were then washed out and incubated at 37°C for 20 minutes to resume protein synthesis. Afterwards, the cells were lysated after 0, 1, 2 and 4 hours. Protein levels were detected using a Luminescent Image Analyzer.

6.2.4.2 Estimation of model parameters

To identify representative values of model parameters for trafficking (τ), endocytosis (ED), maturation (μ) and degradation rates (d_{nm} ; d_m), I simulated

the aforescribed experiments by Ficker et al [358], Ke et al. [359] and Li et al. [360], mimicking their experimental conditions.

Optimal model parameters were computed by minimising the difference between experimental and simulated levels of non-mature and total mature hERG, by minimising the Root Mean Square Error (RMSE). This is given by

$$(6.4) \text{ RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i|^2},$$

where $\hat{y}_i$ is a predicted value, y_i is a measured quantity, and N is the number of points. The total RMSE, calculated through the addition of the independent RMSEs for each of the hERG levels and experiments described above, was used as objective function. A total RMSE=0 indicates a perfect fit.

To lessen the dependence of estimated model parameters on their initial input values, I generated a random set of 1000 initial instances for the parameters with the 'lhsdesign' function from the Mathworks toolkit, which produces a Latin hypercube sample matrix of given size [354]. These were used as starting points along with the objective function for the 'fmincon' function from the Mathworks toolkit [354], which searches for the minimum of the specified objective function given constraints, using a steepest descent method. This function was used to find the distributions of optimal model parameters.

Once optimal values for trafficking (τ), endocytosis (ED), maturation (μ) and degradation rates (d_{nm} ; d_m) were obtained, the non-mature protein synthesis rate (α_p) was derived from the steady-state solution of the model, given by:

$$(6.4) [NM]_{ss} = \frac{\alpha_p}{d_{nm} + \mu},$$

$$(6.5) [M]_{ss} = \frac{\mu}{d_m} \cdot \frac{\alpha_p}{d_{nm} + \mu},$$

$$(6.6) [M_{Mem}]_{ss} = \frac{\mu \cdot \tau}{d_m \cdot ED} \cdot \frac{\alpha_p}{d_{nm} + \mu}.$$

For control hERG trafficking conditions, a relative value of $[M_{Mem}]_{ss} = 1$ is assumed, which yields $\alpha_p = (\mu + d_{nm}) \frac{d_m \cdot ED}{\tau \cdot \mu}$. This assumption also allows the

direct use of $[M_{Mem}]$ as a scaling factor of the hERG current in biophysically detailed action potential models, enabling investigations into the role of altered trafficking or its variability in action potential biomarkers, as discussed below.

6.2.5 Sensitivity analysis of hERG trafficking model

To explore the sensitivity of the model to the variability of the parameters, I simulated a population of models by: i) scaling each parameter from 0.5 to 2; ii) applying Latin hypercube sampling of 100 combinations of parameters; iii) using model simulation for 24 hours to reproduce the conditions of *in vitro* experiments.

For further quantification of the results and model dynamical behaviour, the following sets of markers were calculated for each simulated trace: i) half-life of protein levels, $T_{1/2} = \frac{24}{\log_{1/2}(P_{t=24})}$, where $P_{t=24}$ is a given hERG level at time 24 hours; ii) time to maximum level (T_{max}); and iii) maximum level (L_{max}).

6.2.6 Partial correlation analysis

Partial correlation analysis, which allows calculating correlation between two variables while accounting for the effects of the remaining variables, was used to determine the relationship between the calculated set of action potential biomarkers and the parameters of the trafficking model [361].

Partial correlation was used to determine the relationship between the calculated set of biomarkers ($T_{1/2}$, T_{max} , L_{max}) and parameters of the trafficking model [361].

6.2.7 Simulation protocol of action potential models of human iPSC-derived cardiomyocytes coupled to hERG trafficking

I further explored how the time-dependent changes in the mature hERG on the membrane $[M_{Mem}](t)$ would affect cellular electrophysiology. To do this, I used the Paci2018 model of human iPSC-derived cardiomyocytes, kindly provided by Dr Michelangelo Paci [306]. I coupled the hERG trafficking model with membrane electrophysiology by using the value of $[M_{Mem}](t)$ as the dynamic scaling factor of the hERG current (I_{Kr}).

The different dynamic behaviour of both models (action potential changes occurring in milliseconds; effective hERG membrane levels changing in a time course of hours) also allows for the following coupling strategy for separation of time scales. Firstly, the action potential model is run for 800 s (>12 min) to ensure it reaches steady-state conditions, representing the time point $t = 0$ in the simulations. The hERG trafficking model is then solved up to a given time point, and the value of $[M_{Mem}](t)$ is communicated to the action potential model as a scaling factor of its maximum I_{Kr} conductance. The action potential model is then simulated keeping $[M_{Mem}](t)$ constant for 400 s (>6 min; once again, explored effective changes in hERG membrane levels occurring in a time span of hours) to accommodate to the changes in the hERG current with respect to the previous time point considered. This choice of simulation times provided stable action potential biomarkers in all the hERG trafficking conditions considered. The proposed strategy may only exhibit limitations during fast transients in hERG membrane levels (e.g. right after acute pathway inhibition), for which a strong coupling strategy (updating hERG trafficking inside the action potential model) could be considered.

6.2.8 Action potential biomarkers

I used the following ten biomarkers to quantify changes in action potential under arsenic trioxide action: maximum diastolic potential (MDP), Peak, action potential amplitude (APA), maximum upstroke velocity (VMax), action

potential duration at 10, 20, 30, 50, 70 and 90 % of repolarisation (APD_{10} , APD_{20} , APD_{30} , APD_{50} , APD_{70} , APD_{90}) (see Chapter 3 p. 77 for details on their calculation).

Biomarkers were calculated using R (3.4 version [355]). Firstly I detected beating cycles and after calculated the described set of biomarkers for each individual action potential.

6.3 Results

6.3.1 Estimated parameters of the hERG trafficking model

Table 6-4 Identified set of median model parameters obtained by minimising the difference between simulated and experimental in vitro values of hERG levels. μ : maturation rate; d_{nm} : non-mature degradation rate; d_m : mature degradation rate; τ : trafficking; ED : endocytosis.

Parameter	Value
α_p	0.226
d_{nm}	0.058
d_m	0.207
μ	0.216
τ	2.690
ED	2.317

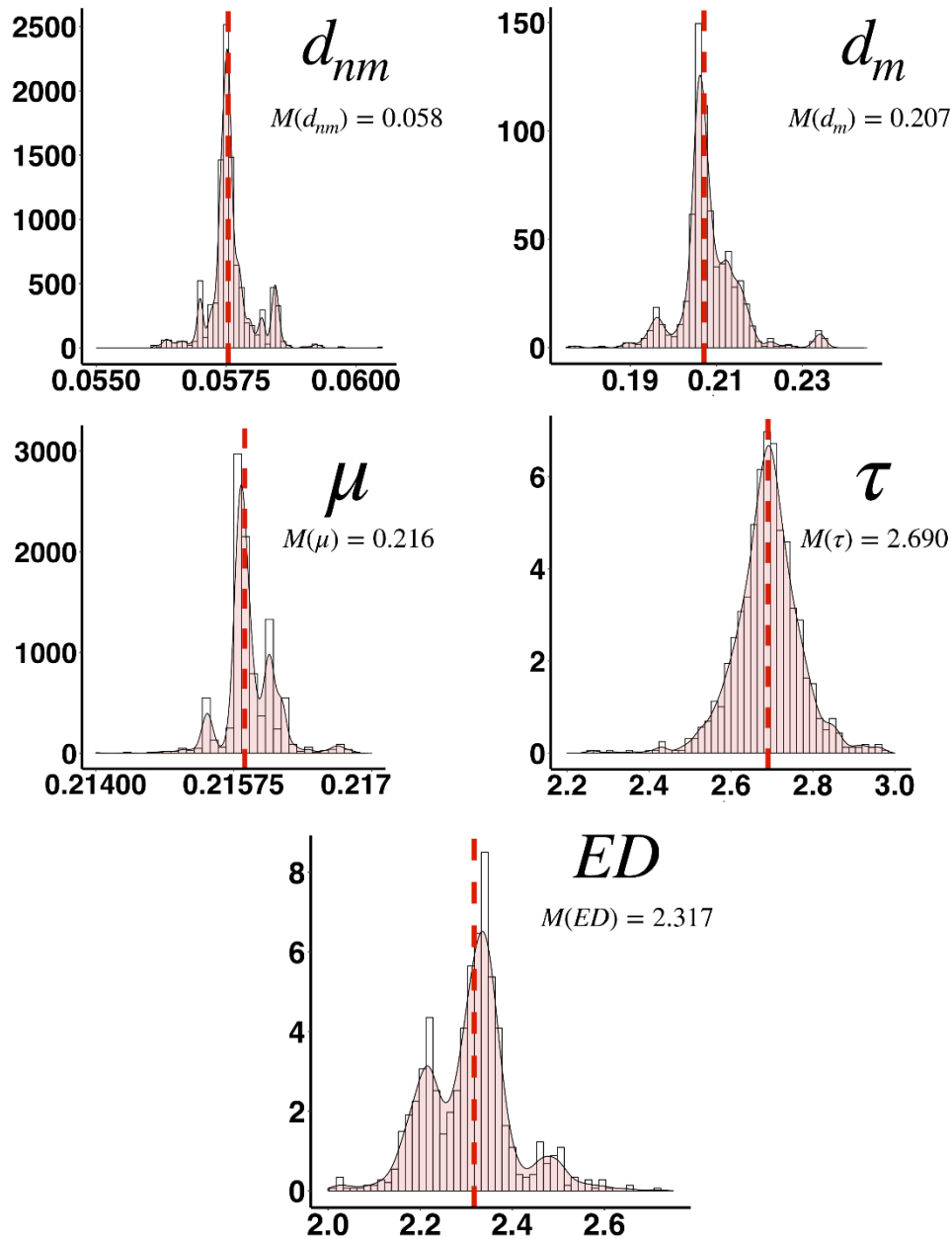


Figure 6-1 Distribution of parameters reached by the optimisation algorithm initialised with 1000 different starting points. For the median of the parameter distributions, histogram plots are overlaid with kernel density curves. μ : maturation rate; d_{nm} : non-mature degradation rate; d_m : mature degradation rate; τ : trafficking; ED : endocytosis.

Table 6-4 and Figure 6-1 show selected parameter values and the distribution of parameters identified by running the optimisation algorithms with the 1000 different starting conditions.

6.3.2 Sensitivity of the hERG trafficking model

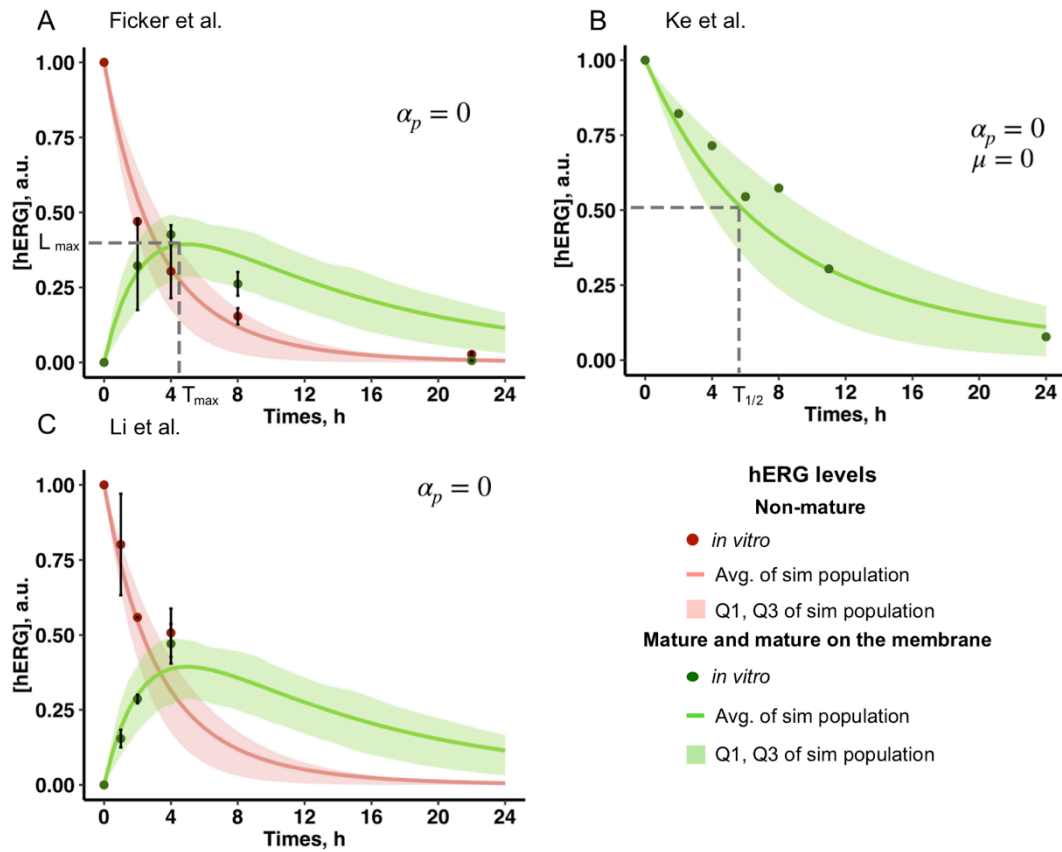


Figure 6-2 Time changes of non-mature and mature hERG. (A) Simulated non-mature and total mature hERG levels (peach and light green, respectively) of a population of hERG trafficking models, and experimental values (dark red and dark green, respectively) reported by Ficker et al [358]. (B) Simulated levels of total mature hERG (light green) of a population of models and experimental values (dark green) reported by Ke et al [359]. (C) Simulated non-mature and total mature hERG levels (peach and light green, respectively) of a population of hERG trafficking models, and experimental levels (dark red and dark green, respectively) reported by Li et al [360]. Avg: average; Q1: 1st quantile (25% of the distribution); Q3: 3rd quantile (75% of the distribution). a.u.: arbitrary units.

Figure 6-2 A-C show average and first/third quantiles of simulated values of non-mature hERG disappearance and mature and mature on the membrane hERG appearance and disappearance, together with experimental data. Values were simulated reproducing experimental conditions ($\alpha_p = 0$), following the sensitivity analysis described in section 6.2.5. It can be seen that simulated values are in agreement with actual experimental *in vitro* data,

showing the same rates of non-mature and mature hERG decay and mature hERG appearance.

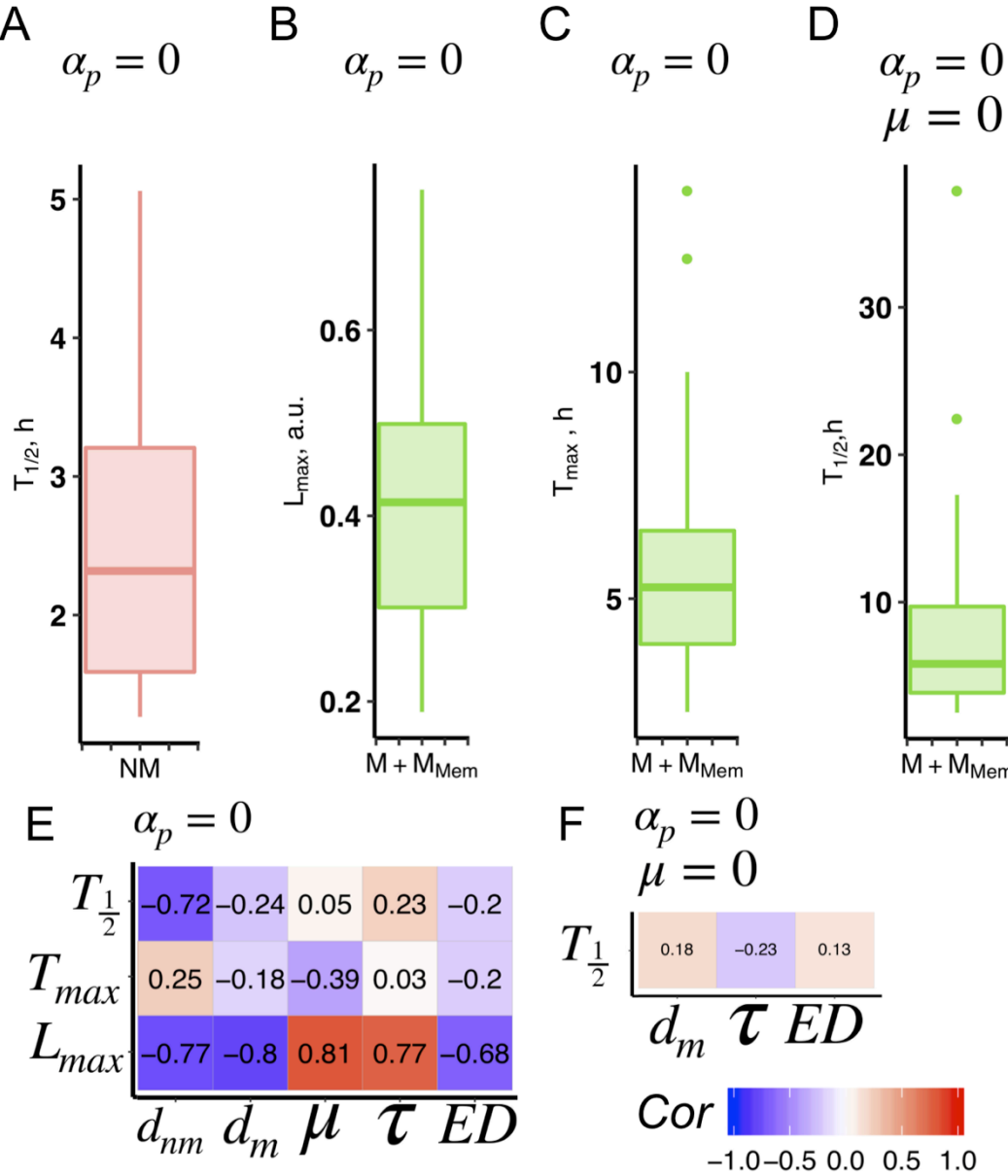


Figure 6-3 (A-D) Box plots illustrating the distribution of the half-life ($T_{1/2}$) of the non-mature (peach) and total mature (light blue) simulated hERG values, time to maximum (T_{max}) and maximum reached levels (L_{max}) for total mature simulated hERG levels (light green). In each box, the central line represents the median, the box limits correspond to the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and the outliers are depicted individually as points. Partial Pearson correlation coefficients between model parameters and hERG biomarkers for (E) non-mature hERG decay ($T_{1/2}$) and total mature hERG production (T_{max} and L_{max}) for the protocol with synthesis blocked ($\alpha_p = 0$), and (F) total mature hERG decay ($T_{1/2}$) for the protocol with synthesis and maturation blocked ($\alpha_p = 0$; $\mu = 0$). [NM]: non-mature hERG; [M]: mature hERG; [M_{Mem}]: mature on the

membrane hERG; a.u.: arbitrary units. μ : maturation rate; d_{nm} : non-mature degradation rate; d_m : mature degradation rate; τ : trafficking; ED : endocytosis.

For further quantification of the results, I calculated half-life values of the simulated non-mature and total mature proteins, along with time to maximum and maximum reached levels (Figure 6-3 A-D). Figure 6-3 E-F shows the partial correlation coefficients between model parameters and the calculated set of biomarkers. For each calculated marker multiple parameters showed significant partial correlation. Maturation rate (μ) has the highest impact on both the simulated non-mature and total mature hERG levels, compared to degradation rates (d_{nm} ; d_m), trafficking (τ) and endocytosis (ED). The higher μ , the faster mature and mature on the membrane hERG reach the maximum level (lower T_{max}) and the higher this maximum level (higher L_{max}). The lower d_{nm} values, the faster the disappearance of the non-mature hERG (higher $T_{1/2}$). Notably, in case of maturation block (Figure 6-2 B and Figure 6-3 D, F), the parameter τ has the highest impact on the levels of mature and mature on the membrane hERG (Cor of -0.23). At the same time, the degradation rate of mature hERG shows a high negative correlation with the simulated total mature hERG levels for the two protocols considered.

6.3.3 Simulations of impaired hERG trafficking

As previously discussed, a wide range of pharmacological compounds and mutations can cause hERG deficiency by affecting different components of cellular hERG trafficking. Here, I explore how changes in the individual model parameters affect simulated values of hERG levels.

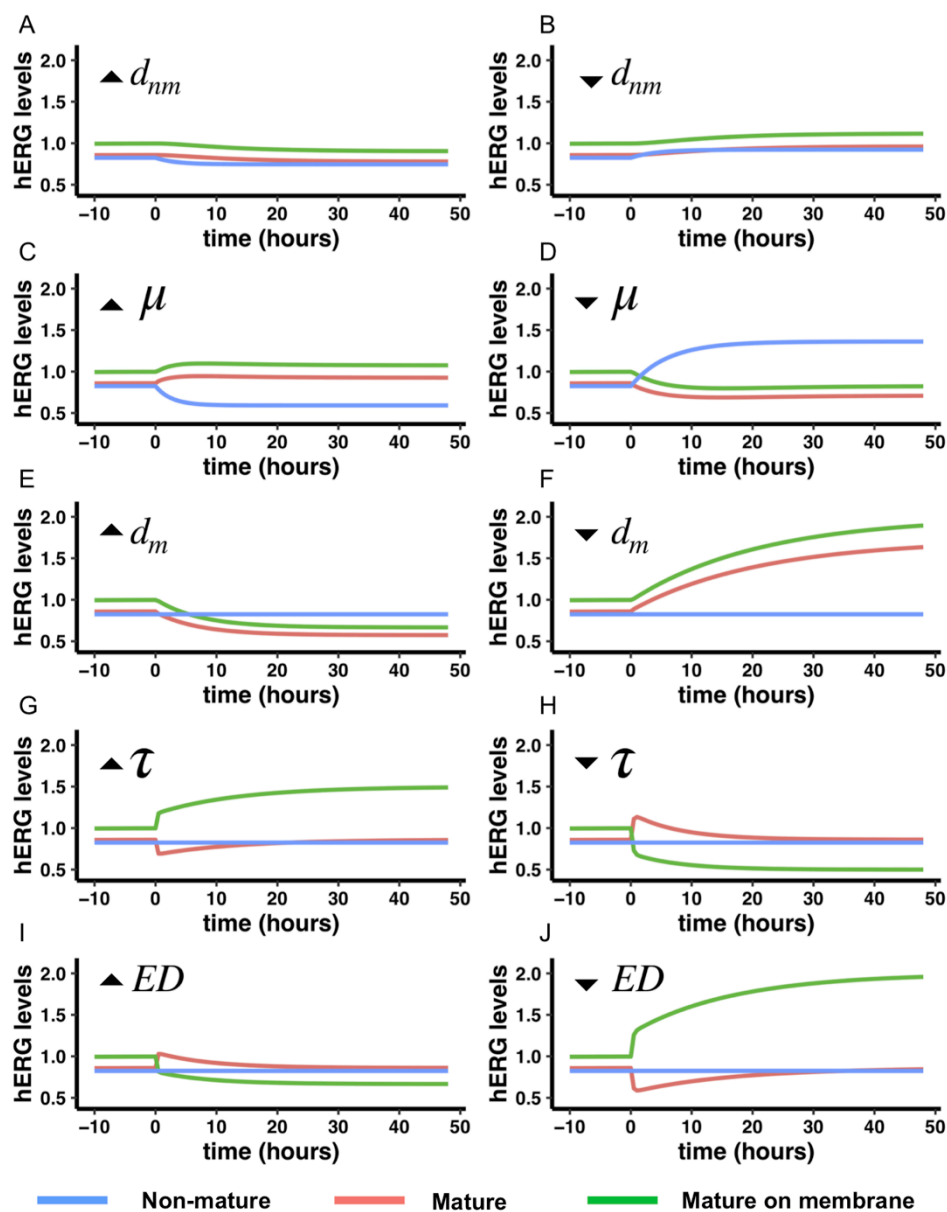


Figure 6-4 Simulations of the time evolution of hERG levels under perturbation of the degradation, maturation and trafficking of the channel. Firstly, the model was initialised with steady-state conditions, and perturbations were introduced at $t=0$ by a 50% increase (left column) and a 50% decrease (right column) in parameters. Levels of non-mature, mature and mature on membrane hERG are displayed in blue, pink and green colours respectively. μ : maturation rate; d_{nm} : non-mature degradation rate; d_m : mature degradation rate; τ : trafficking; ED : endocytosis.

Table 6-5 Simulated levels of mature hERG on the membrane (M_{Mem}) at $t = 48$ hours upon decreased hERG maturation and increased hERG endocytosis.

Scaling	M_{Mem} with $\downarrow \mu$	M_{Mem} with $\uparrow ED$
Control	1	1
25%	0.933	0.833
50%	0.825	0.666
75%	0.611	0.571

Figure 6-4 and Table 6-5 presents changes to the simulated levels of the non-mature, mature and mature on the membrane hERG upon independent increase/decrease of μ , d_{nm} , d_m , τ and ED model parameters right after the start of the experiment ($t=0$). A decreased trafficking has the highest impact on hERG levels on the membrane, showing a 50% reduction of the mature on the membrane hERG after 48 hours (Figure 6-4 H). At the same time, both decreased mature degradation and increased endocytosis lead to a 33% decrease (Figure 6-4 I, E). Alternatively, a two-fold decrease in maturation only results in an 18% decrease of the hERG membrane levels (Figure 6-4 D), while increased non-mature degradation caused a 9% decrease in the hERG levels on the membrane (Figure 6-4 A).

Following these results, I explored how the time-dependent changes in mature hERG density on the membrane would affect cellular electrophysiology. To do this, and as described in section 6.2.6, I used the Paci2018 model of human iPSC-derived cardiomyocytes [306]. I concentrated on two cases of particular relevance: decreased maturation (by 25%, 50% and 75%) and increased endocytosis (by 25%, 50% and 75%), which both resulted in significant reductions of mature hERG on the membrane at long time scales (values after 48 hours listed in Table 6-5).

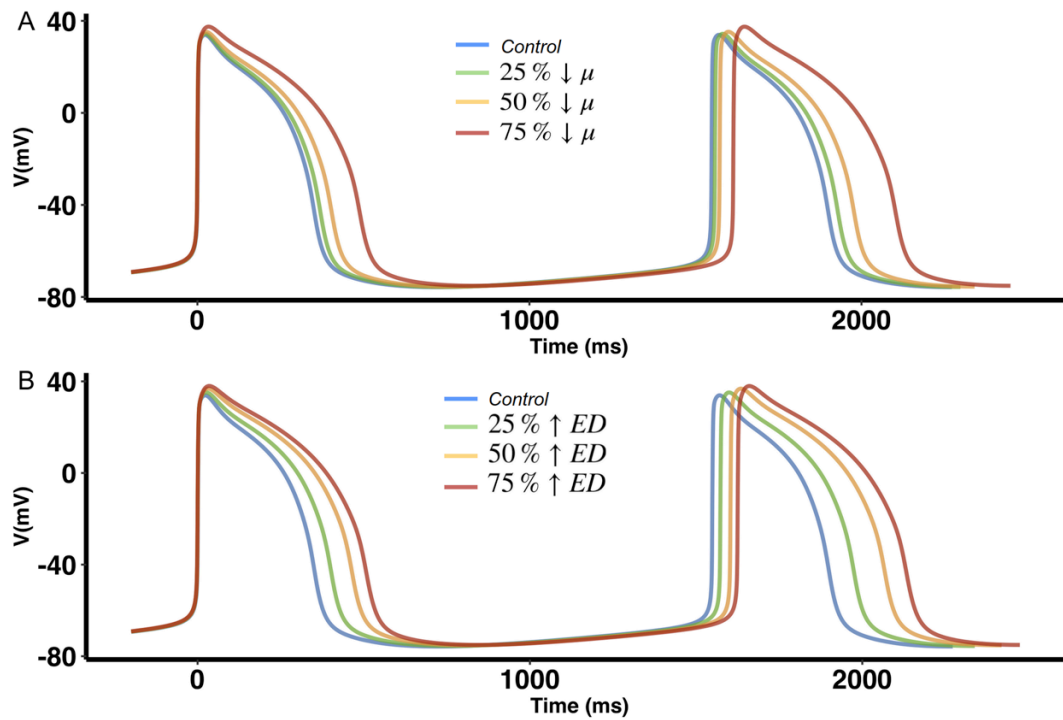


Figure 6-5 Changes in action potentials at $t = 48$ hours upon (A) decreased maturation ($\downarrow \mu$) and (B) decreased membrane stability ($\uparrow ED$) of hERG. The plots show the last two action potentials of the simulated series, aligned at $t = 0$ for improved visualisation of the action potential prolongation changes.

Table 6-6 Quantification of effects on action potential biomarkers at $t = 48$ hours upon simulated inhibition of maturation ($\downarrow \mu$) and decreased membrane stability ($\uparrow ED$) of the hERG channel. Biomarkers were calculated for the last action potential of the simulated series. VMax: maximum upstroke velocity; Peak: peak membrane potential; MDP: minimum diastolic potential; APA: action potential amplitude; APD: action potential duration.

Conditions	VMax, V/s	Peak, mV	MDP, mV	APA, mV	APD90, ms	ADP70, ms	ADP50, ms	ADP30, ms	ADP20, ms	APD10, ms
Control	32	34	-76	110	392	350	319	255	191	96
25 % $\downarrow \mu$	33	34	-76	110	412	370	337	269	203	103
50 % $\downarrow \mu$	33	35	-76	111	446	403	368	293	223	116
75 % $\downarrow \mu$	32	37	-75	113	528	486	444	352	271	150
25% $\uparrow ED$	32	35	-76	111	443	401	366	292	222	115
50% $\uparrow ED$	32	37	-75	112	505	463	422	336	257	141
75% $\uparrow ED$	32	38	-75	113	547	505	461	365	281	158

Action potential changes resulting from inhibition of hERG maturation and decreased hERG membrane stability (simulated as an increased endocytosis) are presented in Figure 6-5, whereas the associated set of action potential biomarkers to quantify these changes is provided in Table 6-6. Both simulated conditions of increasingly impaired hERG trafficking resulted in significant prolongations of action potential durations and increase of the Peak value (Table 6-6 and Figure 6-5). Specifically, a 75% increase in endocytosis results in a 40% increase in APD₉₀, while a 75% decrease in maturation results in a 35% increase in APD₉₀.

6.4 Discussion

As discussed throughout this thesis, a number of agents are known to interact with the hERG channel in a non-acute way, mainly by inhibiting its biogenesis and trafficking, resulting in a rather long-term hERG deficiency (see p. 23 “hERG trafficking inhibition”). Therefore, novel approaches that allow the quantitative characterisation of membrane hERG density changes based on chronic drug exposure are needed. In this chapter, I have presented a new model of hERG trafficking and a framework for understanding how changes in hERG trafficking affect cardiac cellular electrophysiology.

Only a limited number of previous studies have addressed the modelling of specific components of hERG trafficking, degradation and endocytosis using ODEs and experimental in vitro data. However, these were restricted to certain receptors and molecules, including the epidermal growth factor receptor [362], the erythropoietin receptor [363], or low-density lipoproteins [364]. Those models helped to quantify differences in hERG forward trafficking and the stability on the membrane of free and ligand-bounded receptors, and to characterise their physiological function. In contrast, in this thesis, I propose the first model for the complete biogenesis, trafficking and degradation of hERG at the whole cellular function level, which has different dynamic parameters than the previously modelled processes, and a completely different physiological function.

Calibrated with consistent in vitro data from different groups and experiments, the model can reproduce dynamic changes in the levels of non-mature and mature hERG, as observed in the experimental datasets. Interestingly, non-mature and mature hERG levels were most sensitive to variations in maturation rates. Notably, with introduced variability in parameters, the proposed model predicts that the half-life of the non-mature hERG can vary from 1.25 to 5 hours, and the half-life of the total mature hERG from 2.5 to 10 hours. In vitro experiments also confirm a high variability in reported hERG half-life values, with some studies reporting a half-life of the total hERG (mature and non-mature) of up to 17-18 hours [358].

Simulation of pathological conditions in hERG trafficking, by either reducing the maturation rates or increasing endocytosis, showed prolongation of action potential in human iPSC models. Genetic mutations are one of the most significant factors that can cause deficiency of cardiac ion channels. Importantly, hERG mutations are one of the major monogenic causes of life-threatening conditions, such as type-2 Long QT syndrome (LQTS2). This is characterised by profound delayed repolarisation of cardiomyocytes and an increased risk of fatal arrhythmic events. The vast majority of LQTS2 mutations cause the inhibition of hERG protein synthesis, where a defective copy of the transcript has either a stop codon, which terminates the protein synthesis, or a splice site mutation, causing missplicing and generation of incomplete proteins or RNA decay [365]. Other trafficking mutations include the reduction of folding efficiency or increased degradation [365–367]. Accordingly, Mehta and colleagues [368] demonstrated that hiPSCs with a mutation in the transmembrane segment of hERG have lower levels of total mature hERG. Such cells were also characterised by a prolonged action potential (APD_{90} of 1130 ms), compared to the control (APD_{90} of 600 ms).

While decreased maturation or increased endocytosis results in the reduction of the mature hERG on the membrane and the subsequent prolongation of action potential duration and cycle length, previous studies suggest that pharmacological agents can affect multiple stages of trafficking, as in the case

of HDAC inhibitors, which affect both the transcription and acetylation levels of proteins involved in trafficking [45].

6.5 Conclusions

In this chapter, I have described an experimentally calibrated trafficking model of hERG. I have explored how basic changes in model parameters would affect the cellular electrophysiology in hiPSC-CMs, yet the complexity of drug-channel interactions requires a more detailed understanding of their specific modes of action.

In the next chapter, I will explore how the proposed trafficking model, informed by gene and protein expression data and coupled to the action potential model proposed by Paci et al. [306], will be able to reproduce the observed action potential prolongation in hiPSC-CMs caused by chronic inhibition (>20 hours) of hERG trafficking by arsenic trioxide, as an exemplar of application.

7 Simulations of the effects of drug-induced trafficking inhibition on cell electrophysiology

7.1 Introduction

In the previous chapter, I described the proposed trafficking model of hERG and explored how hypothetical inhibition of trafficking processes would affect electrophysiology. In this chapter, I extend its application to investigate hERG trafficking impairment by drug action and its long-term effects in cardiac repolarisation. As mentioned before, a number of agents are known to additionally interact with the hERG channel in a non-acute way, mainly by inhibiting its biogenesis and trafficking [31,32], resulting in rather long-term hERG deficiency. Remarkably, most in vitro studies that investigate drug-induced hERG trafficking inhibition are based on differential total mature protein expression of hERG. However, changes in protein expression may not necessarily translate to changes in channel density on the membrane [369]. Therefore, approaches that allow a quantitative characterisation of membrane hERG density changes upon chronic drug exposure are still needed.

As a proof of concept, I concentrate my investigations in arsenic trioxide. This pharmacological agent is well known to produce negligible acute inhibition of IKr at therapeutic concentrations [196], but can cause Torsades de Pointes after long-term exposure, and is therefore listed in the 'known risk of TdP' category of CredibleMeds (<https://crediblemeds.org/>). Inhibition of hERG maturation has been indicated as a key mechanism of arsenic trioxide, as showed in HEK293 cells stably expressing hERG when chronically incubated with clinically relevant doses [33,127]. A significant increase in action potential duration has also been reported in human iPSC-derived cardiomyocytes chronically exposed to arsenic trioxide [127]. In order to simulate the effects of arsenic trioxide, I first use the gene expression profiles of this drug to identify its mode of action, and hERG protein levels to identify the magnitude of changes in the hERG trafficking model parameters, in order to inform the mechanistic simulation study on its pro-arrhythmic effects presented in this chapter.

7.2 Materials and Methods

7.2.1.1 Characterisation of gene expression profiling experiments

Table 7-1 provides a summary of gene expression datasets analysed in this study. In total four datasets were used to evaluate the response of cells to incubation with arsenic trioxide. Gene expression profiles were collected from the publicly available repositories of the Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo/>).

Table 7-1 Summaries of gene expression datasets for arsenic trioxide. Expression profiles were measured in cells incubated at 37°C. h: hours; N: number.

Agent	Expression datasets	Incubation times, h	Concentration, mM	Cell type	N of Treated samples per case	N of Control samples per (DMSO)
Arsenic trioxide	GSE97460	24	0.4	U87	2	4
	GSE94521	48	1	hESC	5	5
	GSE46909	6	3	Jukat P16 cells	4	4
	GSE14519	6, 24, 48	2	U266 MM1S 8226/S KMS11	1	2

7.2.1.2 Characterisation of in vitro hERG trafficking experiments of drug incubation

Table 7-2 Experimental changes of non-mature hERG ($[NM]$) and total mature hERG ($[M] + [M_{Mem}]$), with degradation rates (d_{nm} and d_m), maturation (μ), trafficking (τ) and endocytosis (ED) active, but protein synthesis blocked ($\alpha_p=0$) upon incubation with arsenic trioxide. Protein levels were measured in HEK293 cells incubated at 37 °C with Western blotting using a radiolabelling system. Data adapted from [33]. SD: standard deviation; n: number of independent experiments; a.u.: arbitrary units; h: hours.

	Control (mean $\pm$ SD, n=3)		As ₂ O ₃ , 3 μ M (mean $\pm$ SD, n=3)	
Times, h	[NM], a.u.	[M]+[M _{Mem}], a.u.	[NM], a.u.	[M]+[M _{Mem}], a.u.
0	1.00 $\pm$ 0.00	0.00 $\pm$ 0.00	1.00 $\pm$ 0.00	0.00 $\pm$ 0.00
2	0.44 $\pm$ 0.16	0.48 $\pm$ 0.18	0.49 $\pm$ 0.30	0.20 $\pm$ 0.04
6	0.22 $\pm$ 0.16	0.50 $\pm$ 0.08	0.09 $\pm$ 0.04	0.27 $\pm$ 0.04

Ficker and colleagues [33] reported the changes of non-mature hERG protein disappearance and mature protein synthesis at time points of 0 to 6 hours using pulse-chase protocols (Table 7-2). They performed monitoring of hERG maturation rate in HEK293 cells incubated at 37°C (3 independent experiments) with arsenic trioxide and without using radioactive labelling. Cells were incubated with [³⁵S]-methionine/cysteine medium for 60 minutes. During the incubation, this radioactive label is incorporated into newly synthesised protein at the 60-minute time point, and can be detected through phosphorimaging. Cells were lysated after 0, 2 and 6 hours to measure changes in the concentration of the labelled proteins, and to detect the disappearance of non-mature hERG and appearance and disappearance of mature hERG upon incubation with arsenic trioxide and under control conditions.

7.2.2 Gene expression and pathway analysis

I performed a differential analysis by comparing control and treated cells for each concentration, incubation type and cell line individually. In total, I performed the differential expression analysis, as described in Section 3.3 (p 66), of 15 such cases using 4 datasets for arsenic trioxide (Table 7-1). The

'limma' package [370] was used to perform differential gene expression analysis for microarray data (GSE94521, GSE46909, GSE14519). The 'Deseq2' package [371] was used to perform gene expression analysis for RNAseq (GSE97460).

Based on the stages of hERG trafficking described in Section 6.2.1 (p 124), I selected the following set of KEGG human pathways related to genetic information processing, and specifically to folding, sorting and degradation (<https://www.genome.jp/kegg/pathway.html#genetic>):

- 1) Unfolded protein response (Hsa04141);
- 2) Folding glycosylation (Hsa04141);
- 3) Endoplasmic reticulum (Hsa04141);
- 4) Vesicular transport (Hsa04130);
- 5) Ubiquitin mediated proteolysis (Hsa04120).

I pre-processed them into a set of eight nodes and edges and used to evaluate the cell response to drug incubation:

- 1) Endoplasmic reticulum (ER) degradation;
- 2) Glycosylation and folding;
- 3) Unfolded protein response;
- 4) Ub-mediated proteolysis;
- 5) Vesicular transport from ER to Golgi apparatus (GA);
- 6) Vesicular transport from GA to membrane;
- 7) Endocytosis.

I used the in silico Pathway Decomposition Network Analysis (iPANDA) [284] algorithm to calculate the pathway activation scores (PAS) for each pathway based on $\log_2FC$ as a metric of differential changes in expression, as described in Section 3.3.2 (p. 68) of this thesis. For these results, I set the term $\sum_j M_{jp} = 0$, in order to analyse only infections of genes included in the pathways. Following [284], the statistical weight coefficients w_i^s were calculated as:

$$(7.1) w_i^s = \begin{cases} 0, & pval_i > p_{max} \\ \frac{\left(\cos\left(\pi \frac{\log pval_i - \log(p_{min})}{\log(p_{max}) - \log(p_{min})}\right) + 1\right)}{2}, & p_{min} < pval_i \leq p_{max}, \\ 1, & pval_i \leq p_{min} \end{cases}$$

where $pval_i$ is the p-value of gene i , and p_{min} and p_{max} are the low- and high-threshold values considered for statistical significance, equal to 0.01 and 0.05, respectively.

For the calculation of the topological weight coefficients w_{ip}^T , I used the 'all_simple_paths' function from the igraph R package [372] to create the graphs from pathway node and edges files, w_{ip}^T values for each pathway p were then calculated by identifying the number of walks that include gene i in this pathway.

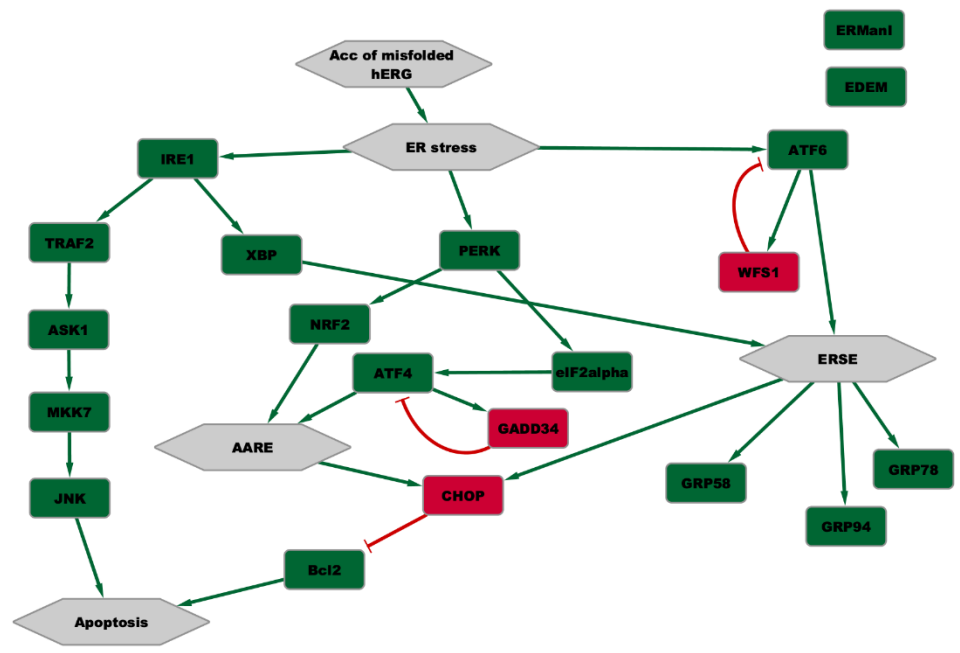


Figure 7-1 Pathway graph showing signalling events in cells in response to the accumulation of misfolded hERG. Each node is a combination of individual genes. Green/red arrows indicate activation/inhibition interactions, respectively. Nodes with positive activation sign ($A_{ip} = 1$) are highlighted in green. Nodes with negative activation sign ($A_{ip} = -1$) are highlighted in red. Acc: accumulation; ER: Endoplasmic reticulum; ERSE: ER chaperons and protein-folding enzymes; AARE: amino acid metabolism Redox/detoxification.

Figure 7-1 depicts interactions and signalling events of the unfolded protein response pathway, which was visualised with Cytoscape software platform [373]. Accumulation of misfolded proteins activates endoplasmic reticulum stress, which could lead to increased production of chaperons and protein-folding enzymes, detoxification and apoptosis.

In order to compare pathways with each other, I normalised each pathway score to its size, which was calculated as $N_{ip} = w_{ip}^T A_{ip}$. The calculated scores are provided in Table 7-3.

Table 7-3 Normalisation scores for each pathway.

Pathway	Normalisation score
ER degradation	9.5
Vesicular transport from ER to GA	8.5
Unfolded protein response	1.5

Glycosylation and folding	11
Ub-mediated proteolysis	43
Vesicular transport from GA to Membrane	12.5
Endocytosis	4

7.2.3 Calibration of the trafficking model

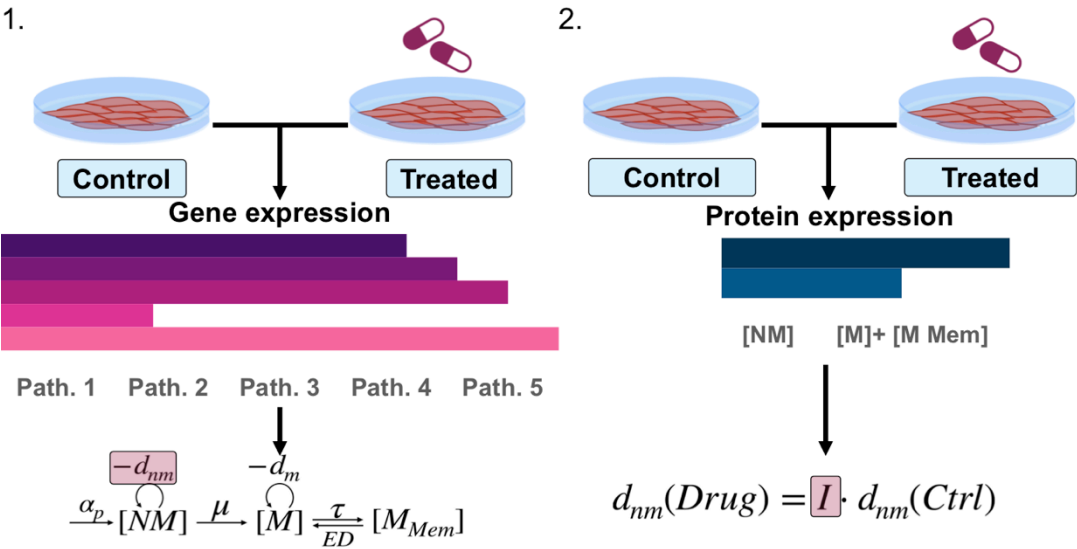


Figure 7-2 Calibration of the trafficking model performed in two steps: 1) identification of parameters perturbed by the drug, and 2) quantification of the magnitude of this perturbation.

Calibration of the trafficking model parameters was performed following a two steps strategy (Figure 7-2). First, gene expression profiles were used to identify signalling pathways perturbed by the drug and their associated model parameters. After, in vitro hERG protein levels were used to quantify the magnitude of these perturbations.

7.2.3.1 Estimation of trafficking model parameters

To identify which parameters of the hERG trafficking model are perturbed by drugs, I used pathway analysis. To quantify their values, I performed a similar optimisation to the one described in the Calibration of the model in Chapter 6 (p 126). I calculated the values of perturbed parameters by simulating the previously described experiments by Ficker et al [33] in Section 7.2.1.2, reproducing their experimental conditions.

The optimal values for perturbed model parameters were computed as in Chapter 6 by minimising the difference between experimental and simulated levels of non-mature and total mature hERG, using the Root Mean Square Error (RMSE). This is given by:

$$(7.2) \text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i|^2},$$

where $\hat{y}_i = \frac{\hat{y}_i(Tr)}{\hat{y}_i(C)}$ is a ratio of predicted treated over control value, $y_i = \frac{y_i(Tr)}{y_i(C)}$ is a measured ratio and N is the number of points. The total RMSE, calculated through the addition of the independent RMSEs for each of the hERG levels and experiments as reported by Ficker et al [33]. A total RMSE=0 indicates a perfect fit. To lessen the dependence of estimated model parameters on their initial input values, I generated a random set of 1000 initial instances for the parameters with the 'lhsdesign' function from the Mathworks toolkit, which produces a Latin hypercube sample matrix of given size [354]. These were used as starting points along with the objective function for 'fmincon' function from the Mathworks toolkit [354], which is searching for the minimum of the specified objective function given constraints. I use it to find the distributions of optimal model parameters.

After I obtained the values of perturbed model parameters for arsenic trioxide, I calculated the inhibition rate as:

$$(7.3) I = \frac{par(Drug)}{par(Ctrl)},$$

where $par(Drug)$ is a specific model parameter after drug application, and $par(Ctrl)$ represents its value under control conditions.

7.2.4 Validation of the trafficking model

7.2.4.1 Characterisation of action potential recordings of drug incubation experiments

Table 7-4 Parameters of APD recorded in hiPS-CMs as reported by Yan et al [127]. Cells were incubated at 37 °C and exposed to 3 µM arsenic trioxide for 24 hours. Afterwards, the

action potential was recorded using the whole-cell patch-clamp technique under current-clamp mode. SD: standard deviation; n: number of independent experiments.

Biomarkers (ms)	Mean $\pm$ SD	
	Control (n = 9)	Arsenic trioxide, 3 μ M (n = 9)
APD ₅₀	588.89 $\pm$ 21.26	959.77 $\pm$ 42.16
APD ₉₀	706.98 $\pm$ 23.71	1164.71 $\pm$ 40.25

Yan and colleagues [127] recorded changes in the action potential of hiPSC-CMs incubated with 3 μ M arsenic trioxide for 24 hours. The whole-cell patch-clamp technique at current-clamp mode was used to measure cells incubated at 37 °C with and without arsenic trioxide. The extracellular solution was a Ca²⁺-containing Tyrode's solution, and the pipette contained 120 mM KCl, 1 mM MgCl₂·6H₂O, 10 mM HEPES, 3 mM Mg-ATP and 10 mM EGTA (pH 7.3 with KOH).

7.2.4.2 Simulation protocol of action potential models of hiPSC-derived cardiomyocytes and coupling with hERG trafficking model

I explored how the time-dependent changes in the mature hERG on the membrane (M_{Mem}) would affect the action potential biomarkers. As described in Chapter 6, I used the Paci2018 model [306] to simulate electrophysiology. I ran it for 800 s (>12 min) to ensure it reaches steady-state conditions, representing the time point $t = 0$ in the simulations. Ionic concentrations were set to the ones used in the in vitro recordings. I specifically set the extracellular concentration of sodium to 136 mM, potassium to 5.4 mM and calcium to 1.8 mM, and intracellular concentration of potassium to 120 mM.

7.2.4.3 Control population of hERG trafficking models

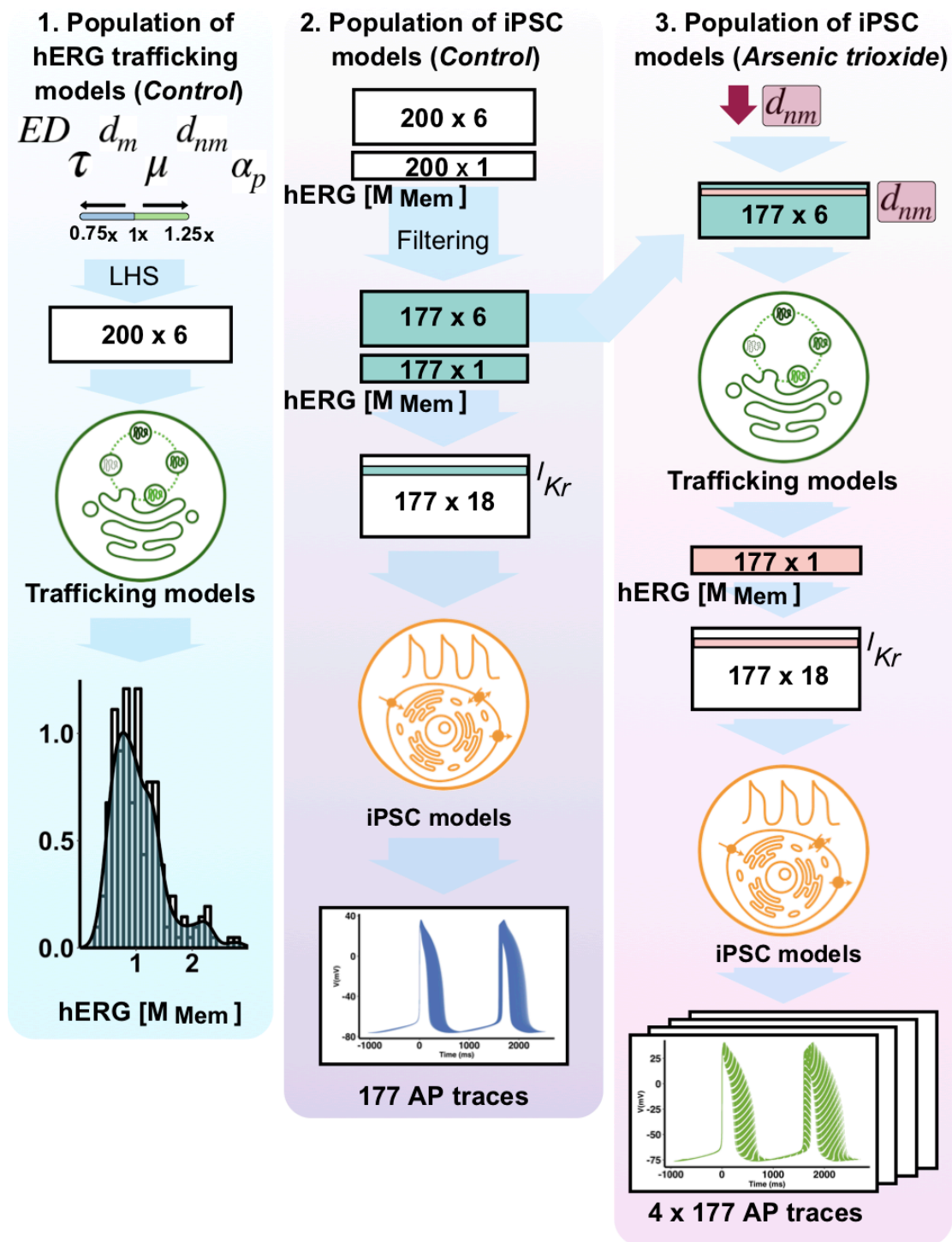


Figure 7-3 Populations of human iPSC models with variability in hERG trafficking properties, under control and arsenic trioxide conditions. Each model set had the same equations, but different hERG trafficking parameters. Firstly, I generated a population of 200 trafficking models. 33 trafficking models were excluded after calibration against physiological values of hERG on the membrane ($[M_{Mem}]$). I then used those 177 hERG values on the membrane values to scale maximum conductance of I_{Kr} and produce 177 human iPSC models.

To construct the population of models used in this study, I explored and compared three different approaches (see Figure 7-3 and Figure 7-4). I first generated 200 trafficking model parameter sets with Latin hypercube sampling (LHS), and simulated values of hERG on the membrane ($[M_{\text{Mem}}]$). Each parameter were sampled in the [75-125]% range of baseline trafficking model values of a_p , d_{nm} , d_m , μ , τ and ED . This parameter set was used to simulate 200 candidate models sharing the same equations but with different trafficking parameters. Afterwards, I filtered the models using IKr maximum conductance values in the [0-200]% range of the original baseline value, which was reported by Paci et al. [305]. This step was done to determine which ones reproduced a behaviour consistent with the experimental data in this previously published study. This first approach is illustrated in detail in Figure 7-3, resulting after calibration in a population of 177 hiPSC models consistent with the experimental data.

7.2.4.4 Control population of human iPSC-derived cardiomyocytes models and coupling with hERG trafficking models

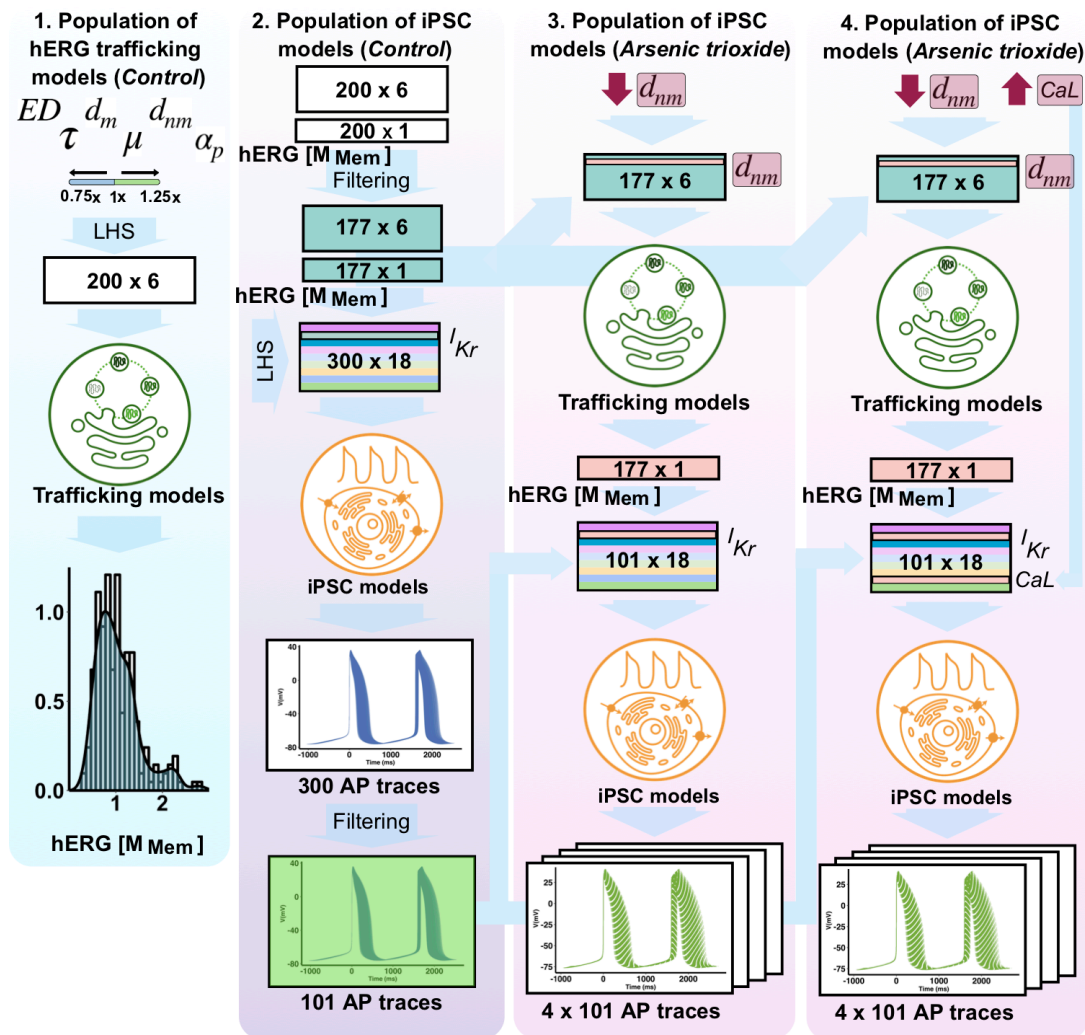


Figure 7-4 Populations of human iPSC models with variability in hERG trafficking and electrophysiology, under control and arsenic trioxide conditions. Each model set has the same equations, but different parameters. Firstly, I generated a population of 200 trafficking models, 33 of which were excluded after calibration with physiological values of hERG on the membrane. I used those 177 hERG membrane values to scale maximum conductance of I_{Kr} , and LHS sampling of conductance and permeability values of the other channels, to produce 300 candidate iPSC models. 101 of the iPSC models and respectively 101 conductance and permeability values were selected after calibration as reproducing control experimental data. The effect of arsenic trioxide on trafficking (panel 3), and on both trafficking and the L-type calcium channel (panel 4), was simulated by scaling the maximum conductance of the I_{Kr} channel according to simulated hERG membrane values, and scaling both I_{Kr} and I_{CaL} , respectively.

I also explored how variation in other currents would impact the results (Figure 7-4). As in Chapter 6, I used the Paci2018 model [306] as a baseline for cellular electrophysiology, and the methodology described by Britton et al. [361] to account for variability. To do this, I generated 300 parameter sets using LHS. Each parameter set contained 12 ionic conductances and permeabilities, as described by Paci et al. [305]. I varied the following currents: fast sodium current (I_{Na}), late sodium current (I_{NaL}), the L-type calcium current (I_{CaL}), the hyperpolarisation-activated cyclic nucleotide-gated funny current (I_f), the inward rectifying potassium current (I_{K1}), slow delayed rectifying potassium currents (I_K s), the transient outward potassium current (I_{to}), the sodium-calcium exchanger (I_{NCX}), the sodium-potassium pump (I_{NaK}) and the calcium sarcolemmal pump (I_{pCa}). Each parameter set was sampled in the [0-200]% range of the baseline values of the conductance and permeability values of those currents. As a result, I simulated 300 models with the same equations for electrophysiology but different ionic conductances and permeabilities.

This candidate population of 300 models was calibrated to select the models in the range with the reported experimental data used by Paci et al. [305]. As per Paci et al. [305], I used Rate, MDP, Peak, APA, V_{Max} , APD_{10} , APD_{20} , APD_{30} , APD_{50} , APD_{70} , and APD_{90} biomarker values to filter the 300 models, calculated as described in Section 3.6.1 (p. 77). The ranges used for calibration of biomarkers are summarised in Table 7-5, resulting in a population of 101 accepted hiPSC models (101 accepted sets of conductances and permeabilities). This population of 101 hiPSC-derived cardiomyocyte models was considered as exhibiting a phenotype in agreement with human experimental data, as measured by the considered set of 10 action potential biomarkers.

Table 7-5 Ranges of biomarkers used to calibrate the candidate population of Paci2018 hiPSC cardiomyocyte models. Values are obtained from Paci et al. [305].

Biomarker (units)	Lower bound	Upper bound
MDP (mV)	-89	-44

Peak (mV)	17	58
APA (mV)	76	139
$V_{\max}$ (V/s)	-27	82
APD ₁₀ (ms)	20	128
APD ₂₀ (ms)	0	290
APD ₃₀ (ms)	59	301
APD ₅₀ (ms)	0	601
APD ₇₀ (ms)	146	631
APD ₉₀ (ms)	1	705

7.2.4.5 Action potential biomarkers of drug action

To determine if the control models reproduced electrophysiological behaviour consistent with the experimental data, I used the following ten biomarkers to quantify changes in action potential under arsenic trioxide action: MDP, Peak, APA, VMax, APD₁₀, APD₂₀, APD₃₀, APD₅₀, APD₇₀, APD₉₀ (See Chapter 3 p. 77). These were calculated at different time points of up to 72 hours, in order to quantify time-dependent changes in action potential upon simulated arsenic trioxide action.

7.2.5 Detection of repolarisation abnormalities

Early afterdepolarisations are abnormal depolarisations that interrupt action potential during the plateau or repolarisation phases of the action potential and are associated with drug action. To detect early afterdepolarisations, I first identified the time of the upstroke in a cycle and then calculated the MDP biomarker for this cycle. Models displaying a Peak higher 0 mV and an MDP greater -44 mV were classified as showing early afterdepolarisations. Models displaying a Peak less than 0 mV were classified as depolarisation failure.

7.2.6 Partial correlation analysis

Partial correlation analysis, which allows calculating correlation between two variables while accounting for the effects of the remaining variables, was used to determine the relationship between the calculated set of action potential biomarkers and the parameters of the trafficking model [361].

7.2.7 L-type calcium current activation by arsenic trioxide

In addition to the effects of arsenic trioxide identified by gene and protein expression in hERG trafficking, previous experimental evidence indicates a role of this drug in upregulating the L-type calcium current [127]. In order to account for these effects, I scaled the maximum conductance of the L-type calcium current according to the reported 95% increase observed in neonatal rat ventricular myocytes [127]. This was modelled on top of the effects of arsenic trioxide on hERG trafficking, to mechanistically investigate the pro-arrhythmic consequences of these two concurrent components of drug action.

7.3 Results

7.3.1 Gene expression and pathway analysis of cell responses to incubation with arsenic trioxide

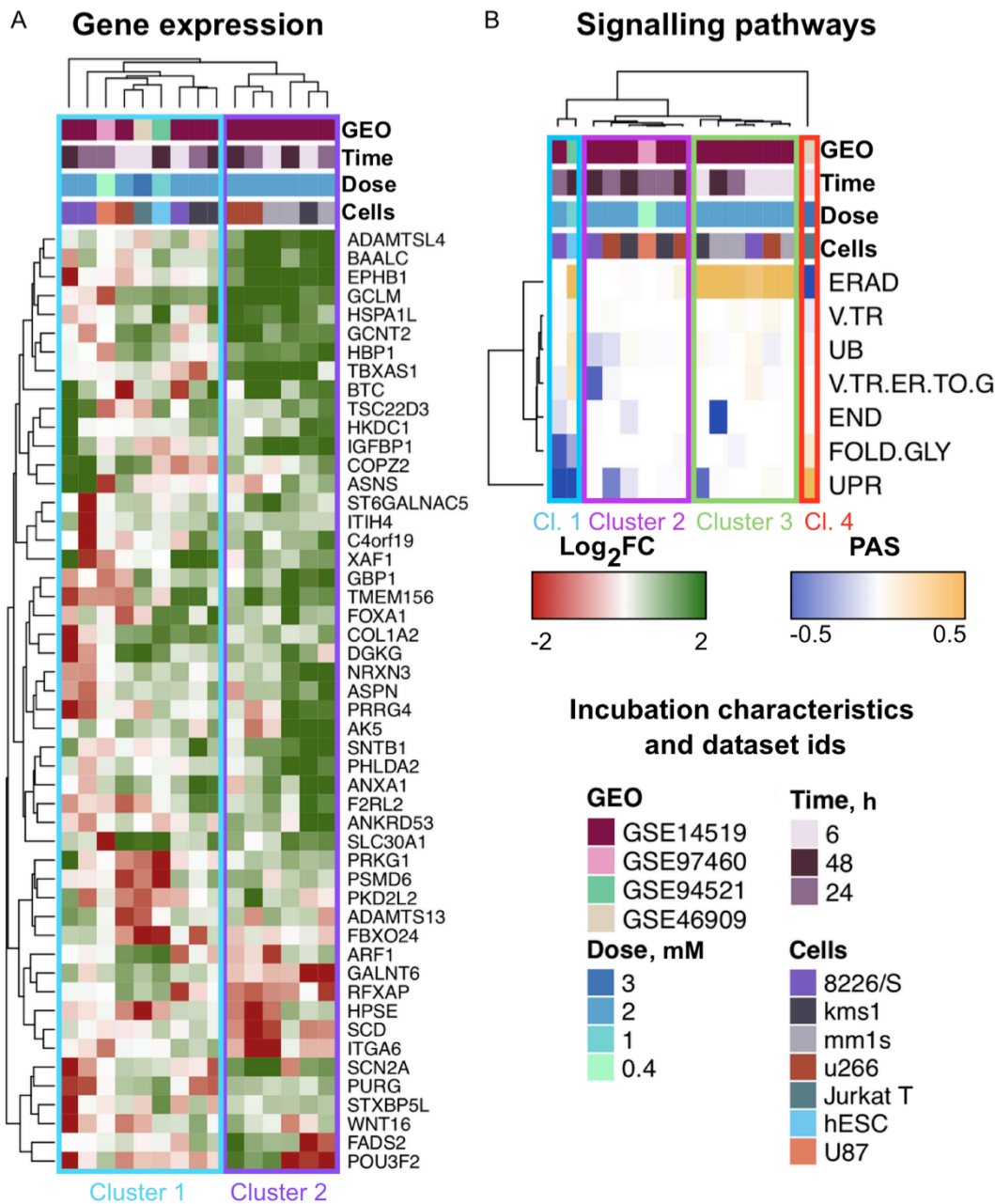


Figure 7-5 Gene expression analysis results of incubation of HEK293 cells with arsenic trioxide. (A) Log₂-transformed fold changes (Log₂FC) of the values of the top 50 differentially expressed genes in cells incubated with arsenic trioxide compared to control groups. Red and green colours denote downregulation and upregulation, respectively. (B) Pathway activation score (PAS) values were calculated for cells incubated with arsenic trioxide compared to control groups. Blue and orange colours denote downregulation and upregulation,

respectively. In both panels, Euclidean distance was used to cluster samples and pathways, and the saturation of the colour represents perturbation amplitude. Cl: cluster; h: hours;

Table 7-6 Pathway Activation Scores (PAS) were calculated for transcriptional profiles of cells incubated with different concentrations of arsenic trioxide. Values are described as median $\pm$ standard deviation. Ub: ubiquitin; ER: endoplasmic reticulum; GA: Golgi apparatus. h: hours; a.u. arbitrary units.

Pathway name	PAS, a.u.		
	6	24	48
ER degradation	0,55 $\pm$ 0,77	0,00 $\pm$ 0,22	0,05 $\pm$ 0,33
Vesicular transport from ER to GA	0,00 $\pm$ 0,03	0,00 $\pm$ 0,02	0,00 $\pm$ 0,20
Unfolded protein response	0,00 $\pm$ 1,52	-0,00 $\pm$ 1,21	-0,00 $\pm$ 1,17
Ub-mediated proteolysis	0,01 $\pm$ 0,04	-0,00 $\pm$ 0,03	-0,00 $\pm$ 0,12
Glycosylation and folding	0,00 $\pm$ 0,06	0,00 $\pm$ 0,20	0,00 $\pm$ 0,10
Vesicular transport from GA to Membrane	0,00 $\pm$ 0,04	0,00 $\pm$ 0,01	0,00 $\pm$ 0,05
Endocytosis	0,00 $\pm$ 0,01	0,00 $\pm$ 0,04	-0,01 $\pm$ 0,23

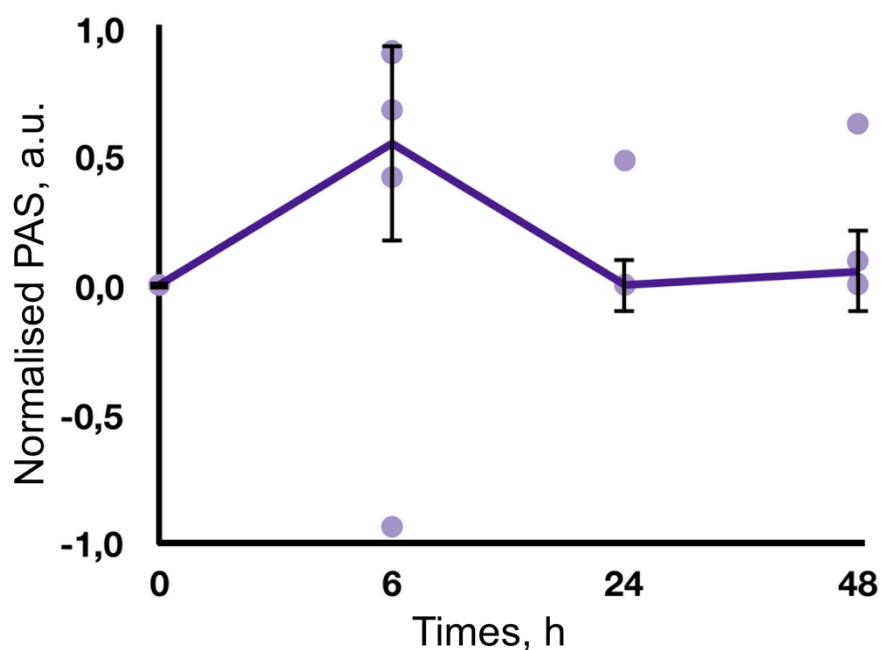


Figure 7-6 Pathway activation scores (PAS) for ER degradation caused by arsenic trioxide. Solid lines represent normalised median values and error bars shows standard deviation values. a.u.: arbitrary units; h: hours.

Figure 7-5 A displays the top 50 genes that change their expression in response to the incubation with arsenic trioxide. An Euclidean distance clustering shows that differential gene expression profiles of cells treated with arsenic trioxide are rather grouped by cell types, forming two large clusters, where Cluster 2 was represented by two cell lines – u266 and mm1s. The calculated PAS for different concentrations and incubation times of arsenic trioxide are summarised and visualised in Figure 7-5 B. In contrast with differential gene expression profiles, Euclidean distance clustering groups PAS values by incubation times instead, forming four clusters, where Cluster 3 was mainly represented by 6 hours incubation and Cluster 2 by 24 and 48 hours. Arsenic-trioxide-treated cells show dysregulation of pathways involved in protein trafficking and, in particular, in the maturation of the hERG protein. More specifically, arsenic trioxide causes the upregulation of gene-encoding proteins that regulate the protein degradation in the endoplasmic reticulum (“ER degradation” pathway;

Table 7-6). Interestingly, and as further visualised in Figure 7-6, there exists an upregulation of the ER degradation pathway during the initial response to drug incubation (time point of 6 hours), remaining stable (zero PAS values) after longer exposure times.

7.3.2 Simulation of arsenic trioxide action on hERG protein expression

Using gene expression data, I therefore inferred that arsenic trioxide results in an upregulation of the ER degradation pathway, which is encapsulated by the parameter describing the degradation of non-mature hERG protein, d_{nm} , in the proposed hERG trafficking model.

In order to quantify the change in model parameters, I performed model optimisation using the in vitro data reported by Ficker et al. [33] for arsenic trioxide (Table 7-2), by running the optimisation algorithms with the 1000 different starting conditions as described in Section 7.2.3 (p.149). Following the results presented in Figure 7-6 for the analysis of the ER degradation pathway, arsenic trioxide action was modelled by assuming a constant perturbed value of d_{nm} . Additional time points of gene expression data would have helped to identify whereas the resulting changes in degradation are time-dependent (only information at 6 hours is available).

Table 7-7 Identified set of model parameters obtained by minimising the difference between simulated and actual in vitro values of hERG levels due to drug incubation. Control values were calculated as specified in section 6.3.1 (p. 132), and are provided for comparison.

Parameter	Value
d_{nm} (Arsenic trioxide)	0.203
d_{nm} (Control)	0.058

Table 7-7 shows the identified set of parameter values for arsenic trioxide action. These were identified by running the optimisation algorithm with 1000 different starting conditions, to lessen dependence on initial parameter values. Arsenic trioxide exposure results in a 3.5-fold increase in the estimated d_{nm} parameter value, which was in line with an increase in the corresponding pathway (“ER degradation”, Figure 7-6).

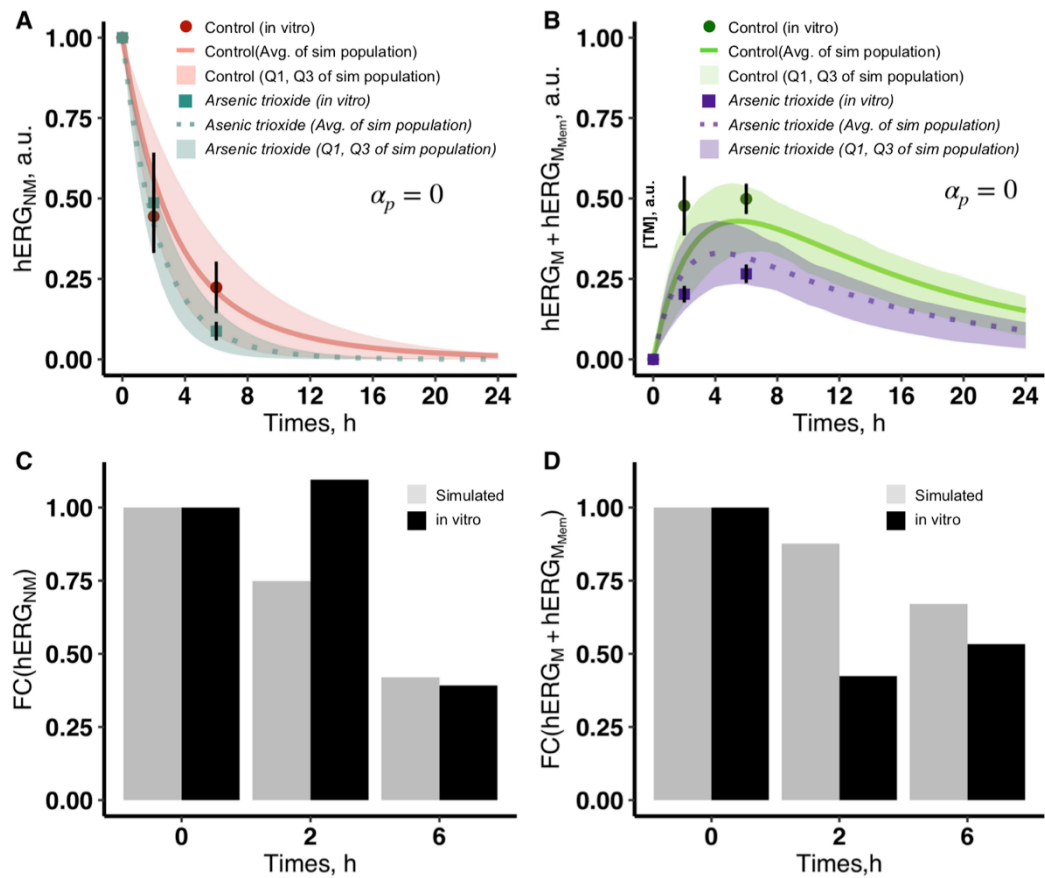


Figure 7-7 Effect of arsenic trioxide on time changes of non-mature and mature hERG. Changes in simulated levels of a population of models of (A) non-mature (peach and light teal for control and arsenic trioxide, respectively) and (B) total mature hERG levels (light green

and light purple for control and arsenic trioxide, respectively) and experimental values of (A) non-mature levels of hERG (dark red and teal for control and arsenic trioxide, respectively) (B) total mature levels of hERG (dark green and dark purple for control and arsenic trioxide, respectively) reported by Ficker et al. [33]. (C) Simulated values (grey) and experimental fold change values of cells treated with arsenic to control levels of non-mature hERG reported by Ficker et al. [33] (black). (D) Simulated values (grey) and experimental fold change values of cells treated with arsenic to control levels of total mature hERG reported by Ficker et al. [33] (black). Avg: average; Q1: 1st quantile (25% of the distribution); Q3: 3rd quantile (75% of the distribution). [NM]: non-mature hERG; [M]: mature hERG; [M_{Mem}]: mature on the membrane hERG; h: hours; a.u.: arbitrary units; FC: fold change treated vs control;

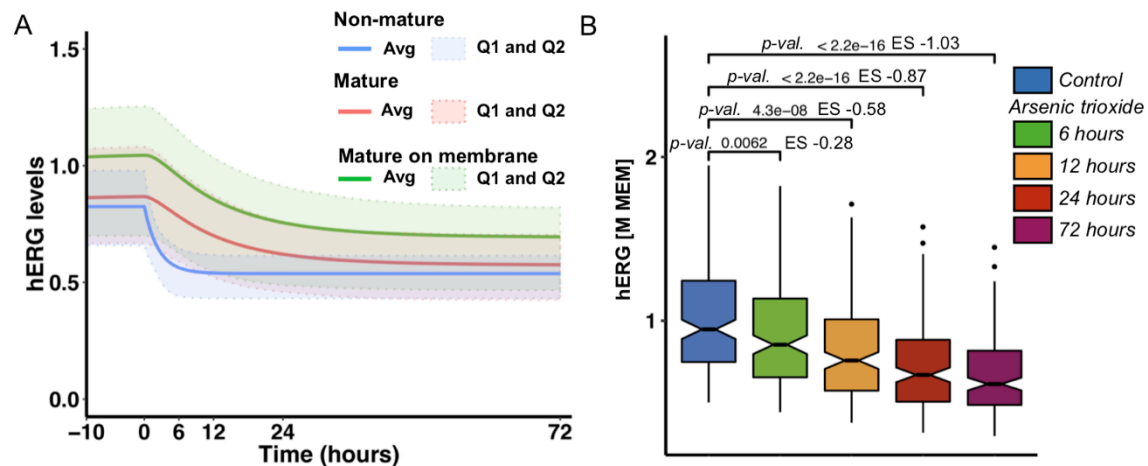


Figure 7-8 Time evolution of hERG levels in the population of models. (A) Average (avg), 1st quantile (Q1, 25% of the distribution) and 3rd quantile (Q3, 75% of the distribution) of changes in hERG non-mature, mature and mature on membrane protein levels upon arsenic trioxide action. (B) Box plot representation of the simulated population of mature on membrane hERG for control and 6, 12, 24 and 72 hours of arsenic trioxide action. In each box, the central line represents the median, the box limits correspond to the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered as outliers, while the outliers are depicted individually as points. P-values are shown for pairwise comparison with control using the Wilcoxon test (non-parametric). P-val: p-values; ES: effective size of change.

Using the parameter values identified above, I first explored how variability in hERG trafficking properties modulate the simulated effect of arsenic trioxide on hERG protein levels compared to the in vitro data reported by Ficker et al. [33]. Figure 7-7 A and B show in vitro data for disappearance of non-mature hERG and appearance of total mature hERG, respectively, together with average and first/third quantiles of simulated values. Figure 7-7 C and D show the ratios of hERG levels treated with arsenic compared to control. These are in good agreement with the reported changes in non-mature and total mature hERG levels after incubation with arsenic trioxide for 6 hours, although may underestimate changes in protein levels during the intermediate phase (time point of 2 hours). Additional time points of gene expression data would help to identify if these differences are due to time-dependent changes in non-mature hERG degradation, or transient activation of other signalling pathway.

Finally, I extended the above-presented results to simulate arsenic trioxide action on the population of hERG trafficking models described in Figure 7-8, under the presence of active synthesis ($\alpha_p \neq 0$). Figure 7-8 A presents changes in the population to the simulated levels of the non-mature, mature and mature on the membrane hERG upon the increase of the non-mature protein degradation associated to arsenic trioxide, right after the start of the experiment ($t=0$). Figure 7-8 B shows the distribution of hERG levels on the membrane for the population of models under control and arsenic trioxide conditions, simulating its hERG trafficking effect after 6, 12, 24 and 72 hours of incubation. On average, 72 hours of incubation with arsenic is estimated to result in a 34% decrease of hERG on the membrane (effective size of 1.03).

Table 7-8 Simulated effect of arsenic trioxide on the mature on the membrane hERG values compared to control conditions in steady state.

hERG	Control (Steady state)	Arsenic trioxide				
		Time, h	2	6	12	24
[NM]	0.825	0.663	0.562	0.541	0.539	0.539
[M]	0.857	0.839	0.770	0.684	0.602	0.567
[M _{Mem}]	0.995	0.981	0.902	0.8	0.701	0.658

The simulation results underpinning Figure 7-8 are further quantified in Table 7-8. Increased non-mature degradation of hERG by arsenic trioxide results in a rapid decrease of the simulated non-mature fraction of the hERG, with a 20%, 32%, 34% and 35% decrease compared to the control values after 2, 6, 12 and 24 hours, respectively. Mature and mature on the membrane hERG values decreased at a slower rate: only a 2% and 1% decrease after 2 hours, a 10% and 9% decrease after 6 hours, a 20% decrease after 12 hours and a 30% decrease after 24 hours. After 72 hours, non-mature hERG levels are decreased by a total of 35%, and mature and membrane-mature hERG levels decreased by 34%.

7.3.3 Simulation of arsenic trioxide action on cellular electrophysiology

To explore how the above described time-dependent changes of arsenic trioxide in membrane hERG levels would translate into changes in cellular electrophysiology, and to validate model with in vitro data reporting action prolongation in hiPSC-CM caused by chronic inhibition (>20h) of trafficking by arsenic trioxide [127], I used the Paci2018 model of hiPSC-derived cardiomyocytes [306].

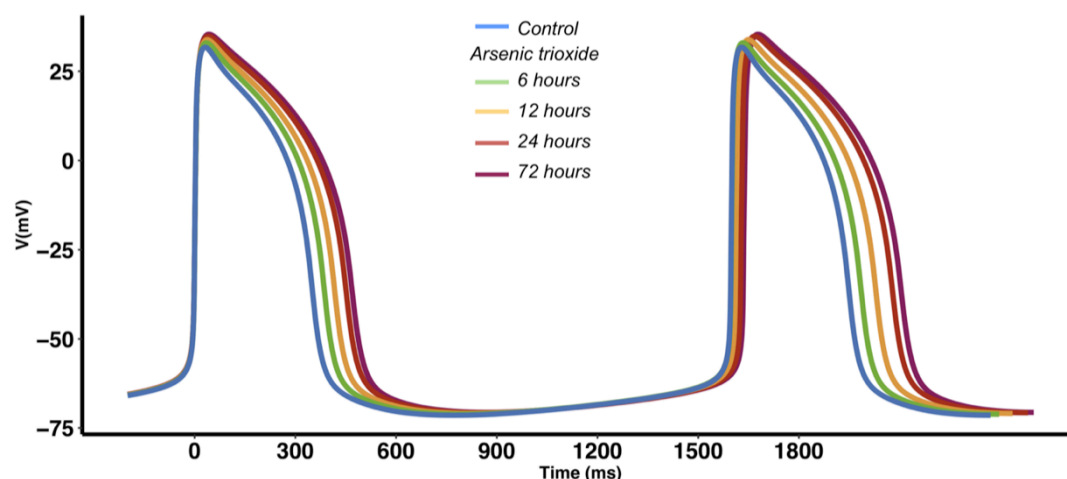


Figure 7-9 Changes in simulated action potential traces of control and the effect of 6, 12, 24 and 72 hours of arsenic trioxide on hERG protein levels for the baseline electrophysiological model. The plots show the last two action potentials of the simulated series of hERG

trafficking inhibition, aligned at $t = 0$ for improved visualisation of the action potential prolongation changes.

Figure 7-9 shows simulated action potential traces in the baseline hiPSC-CM model for control conditions, and the effects of arsenic trioxide after 6, 12, 24 and 72 hours of incubation. In particular, increased degradation of the non-mature hERG caused by arsenic trioxide results in a decrease of VMax and minor increase of APA and peak potential, together with a progressive and marked increase of action potential duration (Figure 7-9 and Table 7-9). Compared to the in vitro results by Yan et al [127] on relative APD prolongation in hiPSC-CM after 24 hours of arsenic trioxide exposure (Table 7-9), both simulated and in vitro APD values display significant prolongation, with a 24% versus 65% increase in APD₉₀, and a 29% versus 63% increase in APD₅₀, respectively.

Table 7-9 Quantitative effects of simulated upregulation of non-mature protein degradation of hERG by arsenic trioxide (baseline model). Biomarkers were calculated for the last action potential of the simulated series. VMax: maximum upstroke velocity; Peak: peak membrane potential; MDP: minimum diastolic potential; APA: action potential amplitude; APD: action potential duration.

Condition	VMax (V/s)	Peak (mV)	MDP (mV)	APA (mV)	APD90 (ms)	ADP70 (ms)	ADP50 (ms)	ADP30 (ms)	ADP20 (ms)	APD10 (ms)
Control	99	32	-71	103	400	355	329	273	214	123
As ₂ O ₃ (6 h)	93	33	-71	104	433	390	363	300	237	140
As ₂ O ₃ (12 h)	90	34	-71	105	462	420	391	323	256	155
As ₂ O ₃ (24 h)	86	35	-71	106	497	453	423	349	277	170
As ₂ O ₃ (72 h)	84	35	-71	106	514	471	440	362	287	178

To extend these results on arsenic trioxide on cellular electrophysiology to the level of a population of hiPSC-CM models and to capture high variability in hiSCP-CM phenotypes, I considered the following three scenarios: (i) arsenic trioxide effects on hERG trafficking, under variability in hERG trafficking properties only; (ii) arsenic trioxide effects on hERG trafficking, under variability in both hERG trafficking properties and rest of ionic currents; (iii) arsenic trioxide effects on hERG trafficking and upregulation of I_{CaL} current, under variability in both hERG trafficking properties and rest of ionic currents. A quantitative comparison of these three scenarios on simulated action potential biomarkers is summarised in Table 7-10, and results presented below.

Table 7-10 Quantitative effects of simulated upregulation of non-mature protein degradation of hERG by arsenic trioxide in populations of models. Biomarkers were calculated for the last action potential of the simulated series. Values represented as mean and standard deviation. VMax: maximum upstroke velocity; Peak: peak membrane potential; MDP: minimum diastolic potential; APA: action potential amplitude; APD: action potential duration; N of EADs: number of models showing early afterdepolarisation; N of DFs: Number of models showing depolarisation failures.

Condition	VMax (V/s)	Peak (mV)	MDP (mV)	APA (mV)	APD90 (ms)	ADP70 (ms)	ADP50 (ms)	ADP30 (ms)	ADP20 (ms)	APD10 (ms)	N of EADs	N of DFs
hERG trafficking variability only												
Control	27±1	32±2	-76±0	108±2	381±80	331±80	300±74	237±60	176±50	87±29	0	0
As ₂ O ₃ (6 h)	26±2	33±2	-76±0	109±2	411±89	363±91	329±85	260±68	196±57	100±35	0	0
As ₂ O ₃ (12 h)	26±2	34±3	-76±0	110±2	441±93	393±96	357±89	282±71	214±59	112±39	0	0
As ₂ O ₃ (24 h)	25±2	35±3	-75±0	110±2	473±94	426±97	387±90	306±70	233±58	125±40	0	0
As ₂ O ₃ (72 h)	25±2	35±3	-75±0	111±2	492±90	445±93	404±86	319±67	244±55	132±39	0	0
Electrophysiological and hERG trafficking variability												
Control	23±16	33±7	-73±3	105±9	325±83	273±68	250±65	199±55	145±46	76±28	0	6
As ₂ O ₃ (6 h)	22±19	36±10	-71±5	107±12	439±184	380±188	346±177	273±144	205±116	121±79	1	6

As ₂ O ₃ (12 h)	22±19	37±10	-71±5	107±12	465±195	406±200	370±187	290±150	219±122	131±84	1	6
As ₂ O ₃ (24 h)	22±19	37±11	-71±6	108±12	495±205	436±211	397±197	310±156	234±127	141±88	1	5
As ₂ O ₃ (72 h)	21±19	38±11	-70±6	108±13	510±212	451±218	411±202	320±159	241±129	147±91	1	6
Electrophysiological and hERG trafficking variability + CaL upregulation												
Control	23±16	33±7	-73±3	105±9	325±83	273±68	250±65	199±55	145±46	76±28	0	6
As ₂ O ₃ (6 h)	23±16	33±7	-73±4	105±9	327±84	276±71	253±66	202±56	147±47	78±28	3	8
As ₂ O ₃ (12 h)	23±14	47±11	-70±8	117±14	581±172	528±179	483±165	369±122	267±96	146±68	3	6
As ₂ O ₃ (24 h)	22±14	48±11	-69±8	118±15	608±177	555±185	507±169	386±122	280±95	154±68	3	6
As ₂ O ₃ (72 h)	21±14	49±11	-69±9	117±16	633±184	580±190	528±171	401±121	292±95	163±69	5	6

7.3.3.1 Arsenic trioxide effects on hERG trafficking under variability in hERG trafficking properties

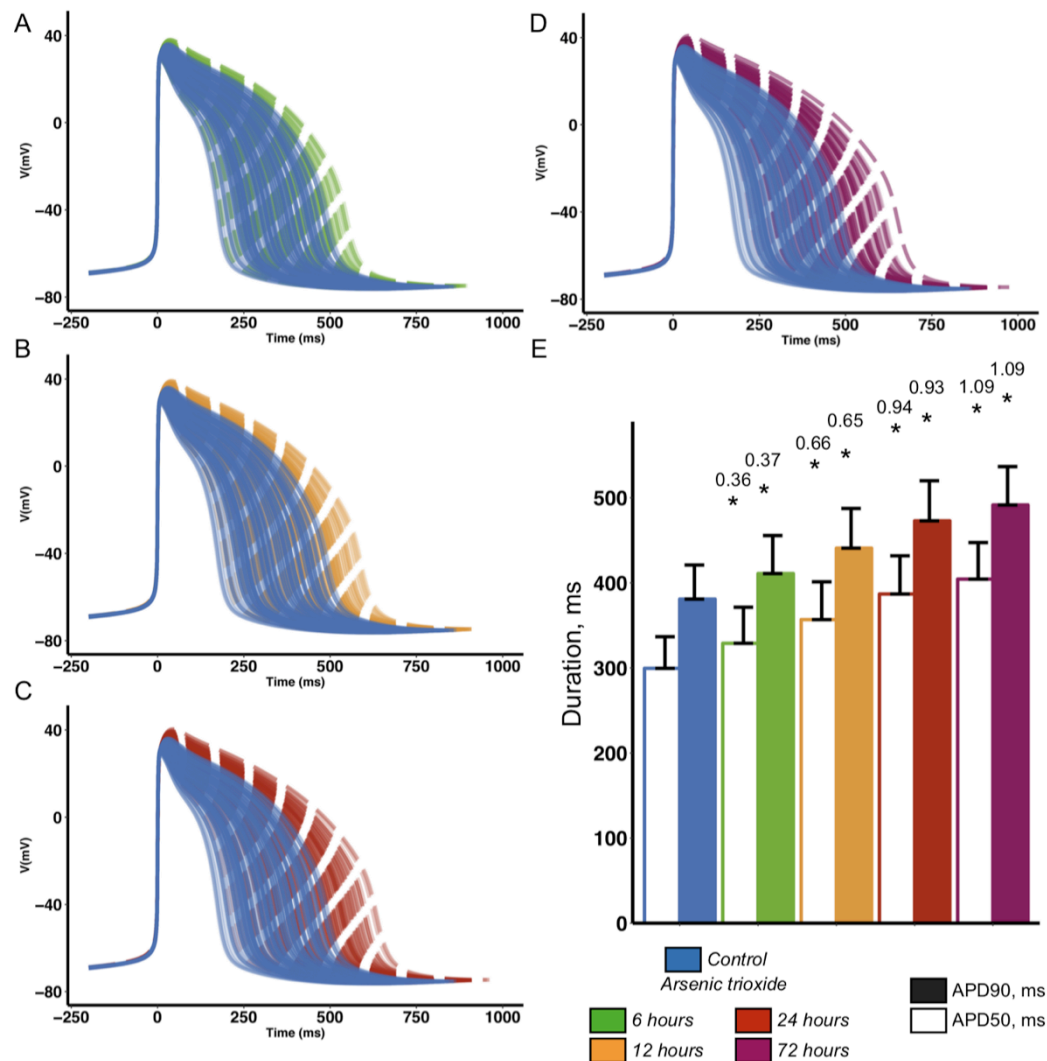


Figure 7-10 Effects of hERG trafficking inhibition by arsenic trioxide on action potential, under variability in hERG trafficking properties only. Simulated action potential traces of (A) 6, (B) 12, (C) 24 and (D) 72 hours of incubation with arsenic trioxide. (E) Statistics of APD₅₀ and APD₉₀, prolonged by arsenic trioxide. The effective sizes of change are shown above their corresponding asterisks, which represent a statistically significant p -value < 0.05 vs control. APD: action potential duration.

Action potential traces for the populations of cells under control conditions and incubated with arsenic trioxide are shown in Figure 7-10 A-D, under considered variability of hERG trafficking properties only. Under this scenario, simulated 72 hours of incubation with arsenic trioxide leads to progressive and significantly prolonged action potential duration, together with a moderate

increase in peak membrane potential (Figure 7-10 and Table 7-10). Results show significant prolongation of action potential duration already after 6 hours of arsenic trioxide exposure, with a mean APD_{90} of the population of cells of 411 ms vs 381 ms in control. Similarly, the 24-hour effect of arsenic trioxide leads to a mean 24% increase of APD_{90} , and a 36% increase of ADP_{50} (Figure 7-10).

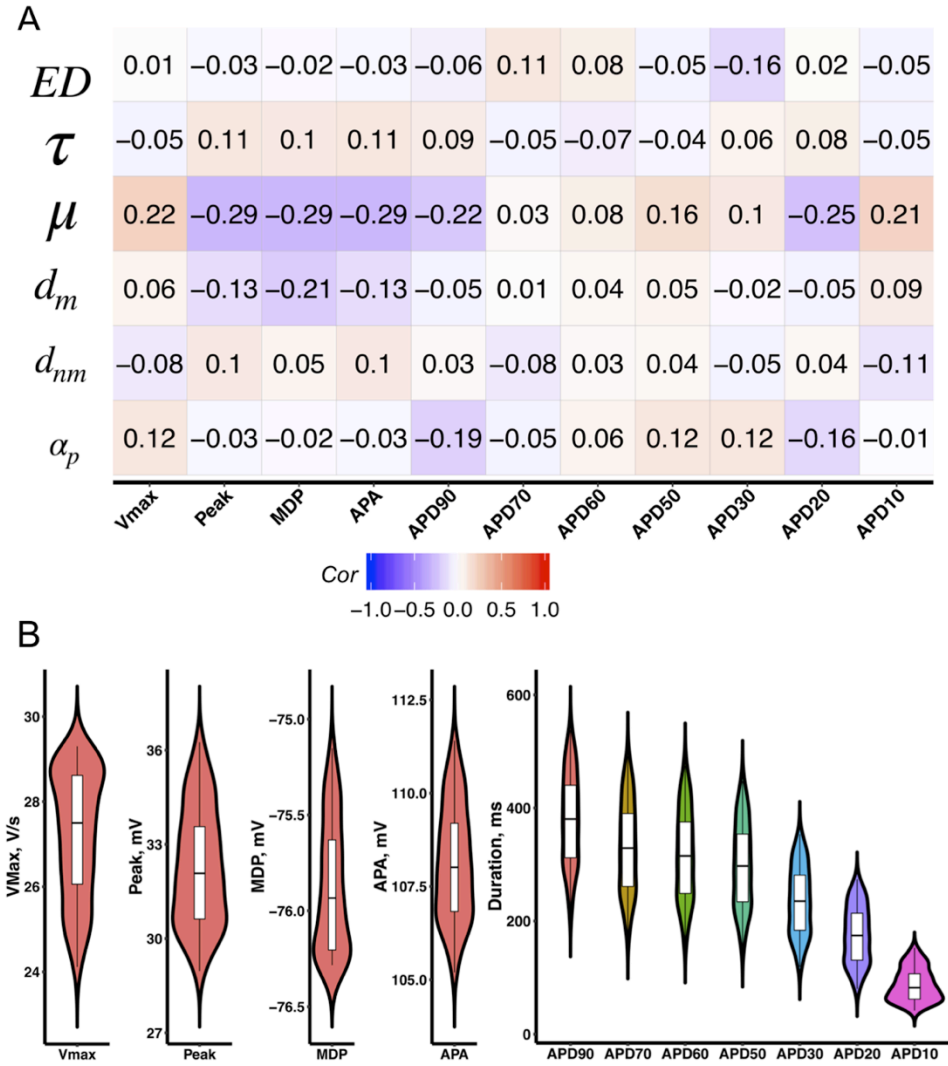


Figure 7-11 Partial correlation analysis in population of hiPSC models with variability in trafficking parameters. (A) Partial Pearson correlation coefficients between action potential biomarkers and trafficking model parameters for the control condition. (B) Violin plots showing the distribution of each biomarker across the population of models. Peak: peak membrane potential; APA: action potential amplitude; VMax: maximum upstroke velocity; MDP: minimum diastolic potential; APD: action potential duration.

In order to determine how the trafficking model parameters contributed to changes in action potential, Figure 7-11 A shows the results of the partial correlation analysis between action potential biomarkers and trafficking model parameters in control conditions. In line with the results at the protein level, the maturation rate (μ) has the highest impact on action potential biomarkers, and in particular the highest negative correlation with APD_{90} . At the same time, the mature degradation rate (d_m) shows a weak negative correlation with MDP, as well as the synthesis rate (α_p) with APD_{90} . Distribution of action

potential biomarkers are shown in Figure 7-11 B. Most of the biomarkers show little variation, with the exception of APD_{90} , APD_{70} , APD_{50} , APD_{30} , and APD_{20} .

7.3.3.2 Arsenic trioxide effects on *hERG* trafficking under variability in *hERG* trafficking and cellular electrophysiology properties

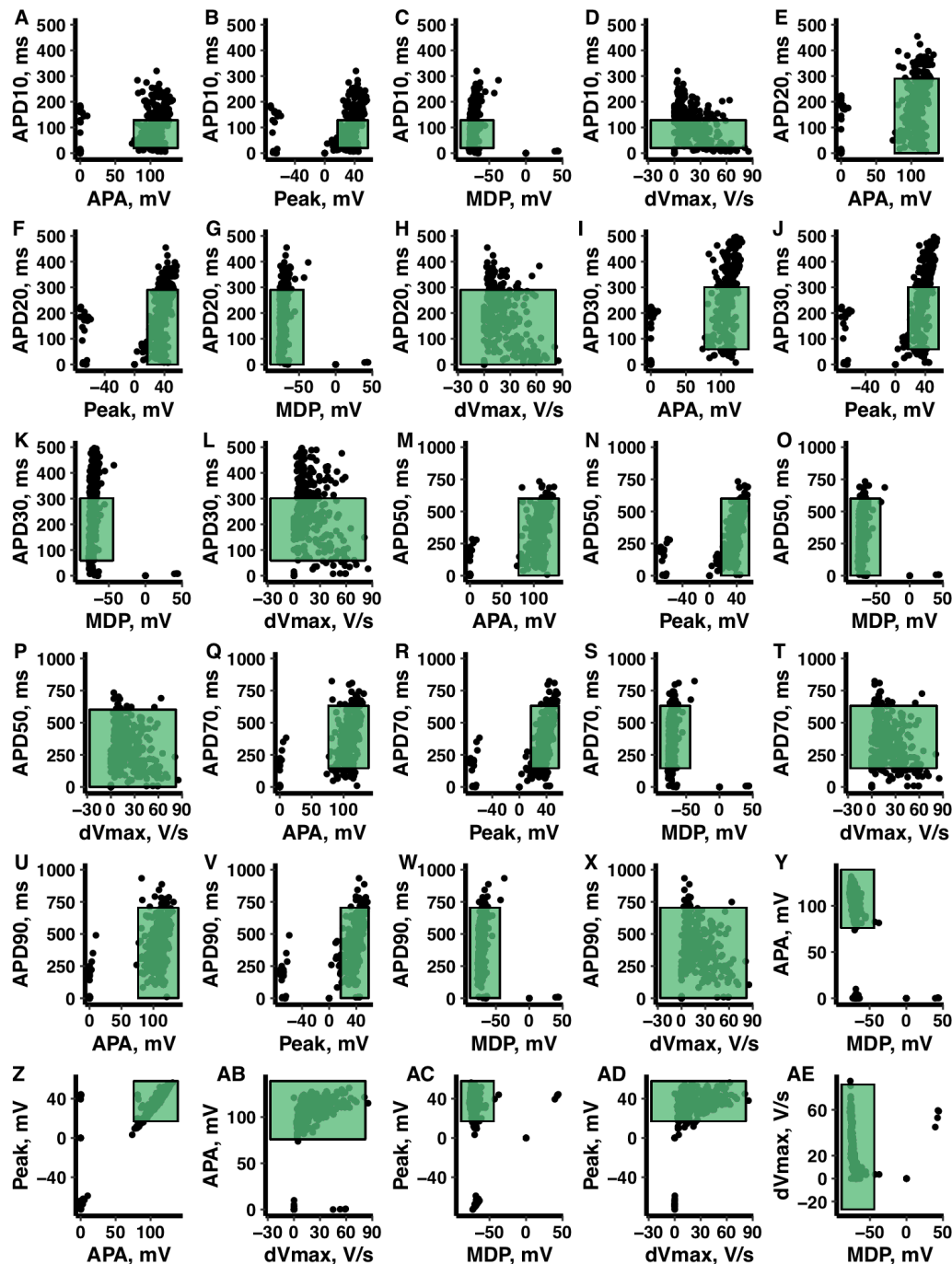


Figure 7-12 (A-AE) Scatter plots showing biomarker values for action potential models accepted in the calibration population (dots in the green rectangles). Green rectangles

represent the lower and upper boundaries for biomarkers obtained from the experimental data [305]. Each plot shows results for a pair of biomarkers. Peak: peak membrane potential; APA: action potential amplitude; VMax: maximum upstroke velocity; MDP: minimum diastolic potential; APD: action potential duration.

Figure 7-12 displays the variability of action potential biomarkers across the initial population of 300 models when variability is also considered in cellular electrophysiological properties, i.e. ionic conductances and permeabilities. Green rectangles illustrate the upper and lower boundaries for each pair of biomarkers considered. In general, the generated population shows both good coverage of the experimental range of markers and overlap with experimental values. After calibration, 101 out of 300 models were accepted. Most of the models were rejected based on the ranges for APD_{10} and APD_{20} biomarkers.

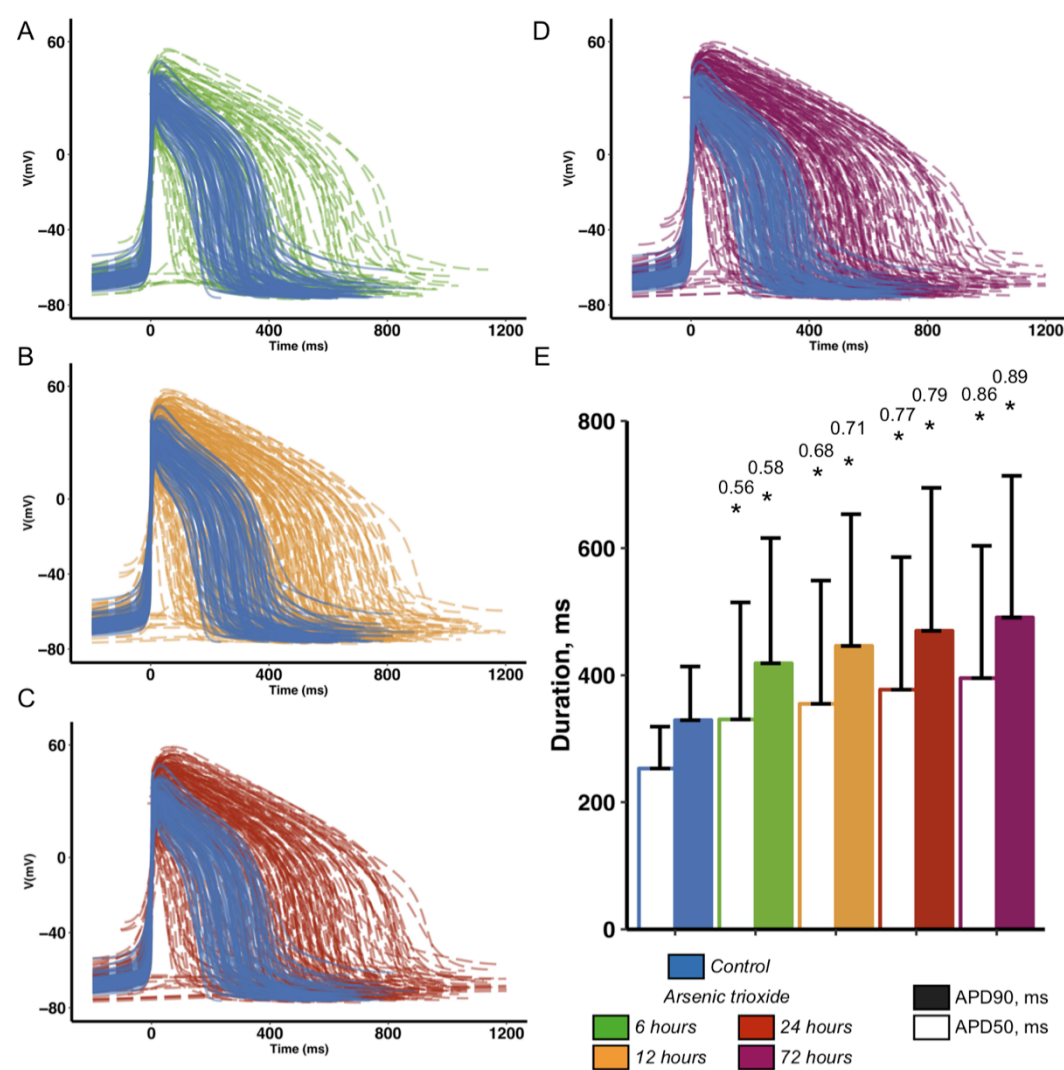


Figure 7-13 Effects of hERG trafficking inhibition by arsenic trioxide on action potential, under variability in both hERG trafficking and cellular electrophysiology properties. Simulated action potential traces of (A) 6, (B) 12, (C) 24 and (D) 72 hours of incubation with arsenic trioxide. (E) Statistics of APD₅₀ and APD₉₀, prolonged by arsenic trioxide. The effective sizes of change are shown above their corresponding asterisks, which represent a statistically significant p-value < 0.05 vs control. APD: action potential duration.

Using this calibrated population of control models representing healthy cells, I first simulated the effect of arsenic trioxide on hERG trafficking. Figure 7-13 shows action potential traces for the simulated control population and arsenic-trioxide-treated models. Results show even more profound prolongation of action potential duration, where the 24-hour effect of arsenic trioxide leads to a mean 52% increase of APD₉₀, and a 59% increase of ADP₅₀ (Table 7-10), in good agreement with the experimental results reported by Yan et al [127]

(Table 7-4). In particular, the time course of action potential prolongation is also markedly less linear than when only variability in hERG trafficking properties is considered. Notably, some of the models failed to produce any action potentials. Those models were excluded from the calculation of biomarkers. In total, 6 out of 101 models showed no action potential. At the same time, one model produced early afterdepolarisations for each simulated incubation time.

7.3.3.3 Arsenic trioxide effects on hERG trafficking and I_{CaL} current under variability in hERG trafficking and cellular electrophysiology properties

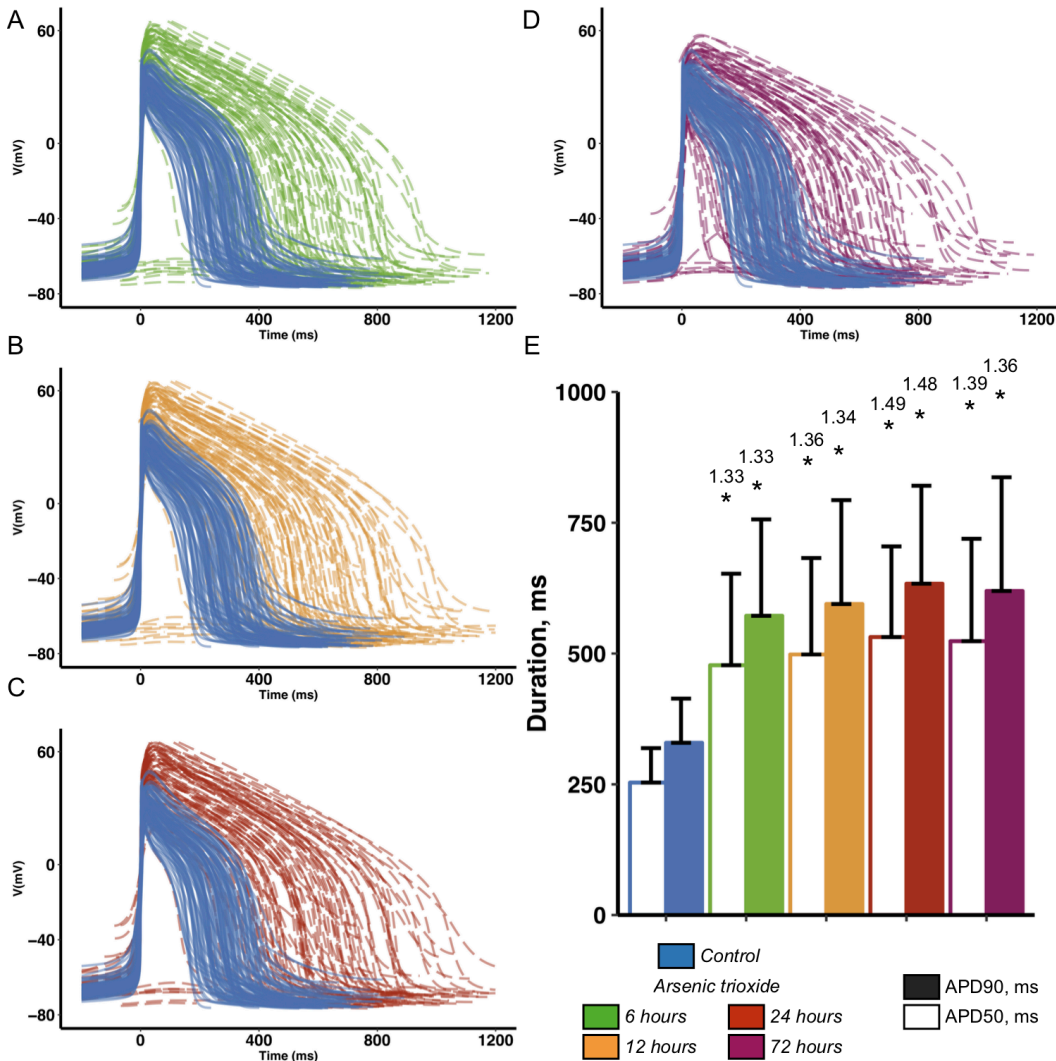


Figure 7-14 Effects of hERG trafficking inhibition and I_{CaL} upregulation by arsenic trioxide on action potential, under variability in both hERG trafficking and cellular electrophysiology

properties. Simulated action potential traces of (A) 6, (B) 12, (C) 24 and (D) 72 hours of incubation with arsenic trioxide. (E) Statistics of APD_{50} and APD_{90} , prolonged by arsenic trioxide. The effective sizes of change are shown above their corresponding asterisks, which represent a statistically significant p -value $< 0,05$ vs control. APD: action potential duration.

Finally, I investigated the effect of arsenic trioxide on hERG trafficking and I_{CaL} upregulation, using the calibrated population of models representing variability in both hERG trafficking and cellular electrophysiology properties. Figure 7-14 shows action potential traces for the simulated control population and arsenic-trioxide-treated models under these conditions. L-type calcium increase by arsenic trioxide resulted in 5% of the models exhibiting afterdepolarisations for the simulated incubation of up to 72 hours. At the same time, 8 out of 101 models failed to produce action potentials for a simulated 6 hours incubation with arsenic trioxide, and 6 for all other incubation times. However, prolongation of action potential duration after incubation for 24-hours resulted in a mean 87% increase of APD_{90} , and a 103% increase of ADP_{50} (Table 7-10), significantly larger than the mean 65% and 63% increase experimentally reported by Yan et al [127], respectively (Table 7-4).

7.4 Discussion

In this chapter, I have presented the modelling and simulation of drug action on hERG trafficking and its impact on cellular electrophysiology of hiPSC cells. Using arsenic trioxide as an exemplar of application, I first identified the set of hERG pathways and trafficking parameters affected by the drug, using gene pathway analysis and calibration with in vitro protein expression data. In addition, I simulated the effect of this drug on the cardiac action potential, both on a single-cell level and on a population of cells. I also compared the inhibition of hERG trafficking induced by arsenic trioxide, secluded and concurrently with the upregulation of L-type calcium current known to be caused by this agent.

Previous studies have extensively investigated whether acute drug effects at the cellular level can be simulated by a combination of single-cell models of cardiac electrophysiology and simple pore-block drug models, the latter informed by measures indicating the potency of drug-channel interactions as described by the Hill equation [290,309,310,313]. Such approaches reproduce experimental data and have proved really useful in simulating disease conditions or variability of acute drug responses within healthy or diseased populations. However, chronic drug effects and trafficking mutations are less explored.

I constructed and validated a mechanistic model of hERG channel trafficking. This model allows the simulation of the amount of mature functional hERG on the cellular membrane as a function of time in order to compare normal versus altered trafficking by drugs, based on their mechanisms of action derived from cellular gene and protein expression profiles. Using this model, I explored the effect of arsenic trioxide, which is known to inhibit trafficking of hERG. This was performed at the protein level, at single-cell level and at the population of cells level.

Folding inhibition by interaction with chaperones has been proposed as a potential mechanism of arsenic trioxide trafficking inhibition [34]. The gene expression analysis conducted in this chapter confirms this mechanism, showing the upregulation of the genes of the endoplasmic reticulum protein degradation pathway, which targets misfolded proteins. Interestingly, the IGFBP1 gene shows upregulation in cells treated with arsenic, suggesting potential ongoing endoplasmic reticulum stress [374].

Increased degradation of non-mature protein results in a 35% decrease of hERG on the membrane in simulations and a subsequent increase of APD, which is also seen in in vitro experiments [127]. Notably, the trafficking model suggests that at 24 hours of incubation with arsenic trioxide, both mature and mature on the membrane hERG levels are still decreasing (Table 7-8).

All four approaches for translating hERG trafficking inhibition by arsenic trioxide considered in this chapter showed significant prolongation of APD. Remarkably, only populations of in silico human iPSC models calibrated using experimental data showed that, under hERG trafficking inhibition caused by arsenic trioxide, physiologically relevant variability in cardiac electrophysiology properties is more prone to produce repolarisation abnormalities, with some models showing early afterdepolarisations, and some even failing to elicit action potentials. Notably, while the effect of longer incubation times was more pronounced on the action potential biomarkers, models already showed repolarisation abnormalities after 6 hours of arsenic trioxide action.

At the same time, concomitant L-type calcium channel upregulation caused by arsenic trioxide resulted in greater prolongation of action potential duration and more repolarisation abnormalities compared to a single 24-hour inhibition of hERG trafficking (93% vs 52% increase in APD_{90} , and 108% vs 58% increase in APD_{50}). Yet, in the in vitro results by Yan et al. in hiPSC-CMs, the action potential biomarkers exhibited a mean 65% increase in APD_{90} and 63% in APD_{50} [127]. This suggest that while it is known that arsenic trioxide activates calcium channels [127], this activation may not be easily translated to human from animal experiments.

7.5 Conclusions

In this chapter and in Chapter 6, I described a new mechanistic model of hERG trafficking and explored the effects of trafficking inhibition on pathological conditions and as caused by a pharmacological agent.

In the next and final chapter, I will present an overview of the work presented in this thesis and summarise the major findings. I will then conclude the thesis with a discussion of the possible future directions of this work.

8 Conclusions

8.1 Summary of the thesis

Adverse cardiovascular effects are one of the most common complications of medications. Such effects vary in severity, ranging from hypertension to sudden cardiac death. In terms of their time course of manifestation, some can develop with an sudden onset, while others develop chronically. The diversity of the underlying mechanisms of drug-induced cardiotoxicity complicates the prediction of cardiac safety.

Due to the variability of the mode of action of drugs on heart tissue and the high heterogeneity of the biological data, the prediction of drug-induced cardiotoxicity is difficult. Computational modelling and simulation has mostly been focused on predicting cardiac events, such as Torsade de Pointes, or on predicting a single type of cardiac complications.

In this thesis, I aim to develop an *in silico* method for prediction of drug-induced cardiotoxicity. To achieve this goal, I have employed two complementary computational methods. I first developed a machine learning approach for predicting a broad range of drug-induced cardiac adverse effects, which goes beyond acute effects of drug action. I then proposed a new mechanistic model of hERG trafficking, which I used to analyse the effects of arsenic trioxide, an antineoplastic agent known to cause Torsade de Pointes without acute block of hERG but through its trafficking inhibition.

In Chapter 2, I began with a review of the known mechanisms of drug-induced cardiotoxicity and of 77 pharmacological agents from different therapeutic areas, including anti-cancer therapies and antivirals. I demonstrated that many of these diverse drug classes share common modes of action on adverse cardiac events, with cardiotoxicity frequently resulting from the simultaneous interruption of key myocardial functions and viabilities.

In Chapter 3, I reviewed the methods of computational analysis of drug-induced cardiac dysfunction. I summarised the current literature on gene expression analysis of cardiotoxic drugs and biomarkers of cardiotoxicity. I then introduced methods of gene expression and signalling pathway analysis. I then provided an overview of machine learning approaches in cardiac safety modelling and described the key challenges of those types of methods. Following this, I described simulations and modelling approaches for cardiac safety and biomarkers used to characterise the outputs of those approaches. Finally, I outlined key databases with potential applications in drug discovery and safety pharmacology, including databases of transcriptional drug profiles and databases of reports of drug cardiac safety in humans.

In Chapter 4, I described my data collection process from the public domain. The collected dataset contains cellular responses to incubation with drugs (transcriptional profiles) and comprehensive information on reported cardiac-adverse reactions induced by those drugs, together with their structural information. I explored the co-occurrence of multiple types of cardiac side effects and their prevalence in specific drug classes. The initial analysis of this dataset showed an association between different cardiac side effects. At the same time, it also revealed an imbalance of the number of safe and drugs with reported cardiac adverse reactions. I also found that the dependencies between cardiac drug safety in the database and their transcriptional profiles and structural information are non-linear.

The data collected in Chapter 4 was used to build a machine learning predictor of drug-induced cardiotoxicity in Chapter 5. I presented a multi-label machine learning classification of the six major forms of drug-induced cardiac side effects, including cardiac disorder signs and symptoms, cardiac arrhythmias, heart failure, coronary artery diseases, pericardial disorders and myocardial disorders. I introduced the chain of classifiers with nested stacking as prediction model for this task, which takes into account target label dependencies. This model demonstrated superior performance in discriminating safe and unsafe drugs compared to models that treat target labels independently. The chapter additionally highlights the importance of a

robust cross-validation strategy and testing set design by comparing leave-drug-out and random cross-validation strategies, and evaluating results in an independent testing set. I also compared the accuracy of cardiotoxicity prediction across different classes of drugs, showing that cardiac arrhythmia-related safety is predicted more accurately for cancer and anti-inflammatory agents than for cardiovascular and central nervous system agents.

In Chapter 6, I have presented a new modelling approach to account for biogenesis, trafficking and degradation of hERG in cellular models of cardiac electrophysiology. I began with the characterisation of the key stages and regulatory proteins of hERG trafficking, which encodes the I_{Kr} channel. I formulated the dynamic model of hERG trafficking in a cell based on protein dynamics, characterising the processes known to mediate the biogenesis and trafficking of hERG and which can be modulated by pharmacological agents. Among other factors, the model considers a non-mature hERG synthesis rate, the maturation rate of hERG, membrane trafficking, mature and non-mature protein degradation rates and surface stability. Model parameters were inferred from in vitro trafficking experiments. Using this model, the amount of mature, functional hERG on the cellular membrane was simulated as a function of time for comparing normal and altered trafficking by different trafficking mechanisms. Afterwards, I analysed how changes in trafficking, like a decreased maturation rate, affect cardiac electrophysiology in hiPSC-CMs action potential models.

Chapter 7 presented the simulation results of drug-induced hERG trafficking impairment. I investigated the effects of arsenic trioxide, known to produce negligible acute inhibition of I_{Kr} at therapeutic concentrations, but to cause Torsade de Pointes through trafficking inhibition. By introducing the amount of mature functional I_{Kr} as a function of time into a population of hiPSC-CMs action potential models, I addressed the exploration of drug effects at different time scales and modelled different scenarios of drug-ion channel interaction. I simulated the effects of arsenic trioxide as a trafficking inhibitor, comparing simulation outputs against hiPSC-CMs incubated with arsenic trioxide for 24 hours. I also explored the contribution of arsenic trioxide in upregulating the L-

type calcium current to prolongation of the action potential duration, showing that hERG trafficking inhibition alone by this drug might be sufficient to recapitulate experimental findings.

8.1.1 Key contributions and main findings

The first major contribution of this thesis is the development of approaches for the modelling, simulation and prediction of drug-induced cardiac reactions that go beyond acute effects. These were described in Chapters 5 and 6, using the data collected and presented in Chapter 4. I proposed different but complementary approaches for modelling with machine learning methods capable of capturing complex dependencies between drug structure and transcriptional profiles, cardiotoxicity and simulations, thus providing insights into the mechanisms of cardiotoxicity. I proposed and implemented a chain of classifiers with a nested stacking approach that classifies drugs by their risk level, has the ability to handle large heterogeneous data, and attains relatively accurate predictions for up to six forms of drug-induced cardiotoxicity.

The second major contribution of this thesis is the presentation of the first-in-its-kind model of hERG trafficking, described in Chapter 6. While existing modelling approaches reproduce experimental data and have proved really useful in simulating disease conditions or variability of acute drug responses within healthy or diseased populations, chronic drug effects and trafficking mutations are less explored. Therefore, the complexity of trafficking disruption requires suitable models to capture these interactions in order to augment the predictive value of *in silico* approaches. This new mechanistic model of hERG trafficking was shown to be easily coupled with existing action potential models of cellular electrophysiology, and was able to reproduce the effects of hERG trafficking impairment under different conditions. The results of this work suggest its potential to complement existing computational methods in general and drug discovery tools in particular.

The final major contribution of this thesis is the study presented in Chapter 7. This work highlights the benefits of using the different computational methods

developed in this thesis for the prediction of drug action, and represents a step forward towards more rational drug discovery. First, the gene expression data collected in Chapter 4 is exploited here to elucidate the mechanisms of arsenic trioxide in hERG trafficking impairment. The hERG trafficking model developed in Chapter 6 was then shown to reproduce the electrophysiological effects of arsenic trioxide in hiPSC-CM action potential models, including action potential prolongation and development of repolarisation abnormalities, as observed in *in vitro* experiments. In particular, the results of this chapter indicate that inhibition of hERG trafficking alone by arsenic trioxide is sufficient to recapitulate the action potential prolongation caused by this drug in hiPSC-CMs. An upregulation of the L-type calcium current has also been proposed as an additional mechanism of action potential prolongation by arsenic trioxide [33,127]. However, the inclusion of this mechanism (as estimated from animal experiments) led to a significant overestimation by almost two folds of the observed action potential prolongation in hiPSC-CM recordings, which may indicate inter-species differences in this mechanism of arsenic trioxide action in human versus experimental animal models.

8.2 Future directions

8.2.1 Incorporating information about related cardiac side effects

The results of Chapters 4 and 5 have shown associations between different cardiotoxicity types, and that it is important to preserve such associations in the modelling of drug cardiac safety. Kidney diseases [375] or metabolic disorders [376] are also well-known complications that are usually associated with heart disease. The current approach and the proposed model can be further extended to incorporate information about other potential side effects known to be caused by the analysed drugs, in order to increase discriminative power and improve its prediction accuracy.

8.2.2 Further experimental validation of the dual machine learning predictor of six cardiotoxicity forms

The dual transcriptomic and structural predictor of cardiotoxicity described in Chapter 5 was built and validated using expression data profiled with microarray technology. However, alternative gene expression platforms, such as bulk or single-cell RNA-seq, are gaining increasing popularity in toxicity studies [377,378]. One of the major advances of the RNA-seq expression profiling is that it offers opportunities for analysis of any transcript, while microarray platforms rely on a predefined set of transcripts, as the L1000 platforms used in this study that measured 971 landmark genes. A next step would be to investigate how a broader range of profiled transcripts, including mRNA and miRNA (which can be measured, for example, with RNA-seq technology) would affect the predictive power of the proposed algorithms.

It should also be noted that the types of chemistry of the explored molecules are somehow limited. Medicinal chemistry, following the increased diversity of biological targets, has also increased the diversity of scaffolds, and of substructures of molecules with known properties, that are now used to design new medications [379]. Therefore, any computational model should be constantly updated with the new types of chemistry of drugs.

8.2.3 Simulation of hERG trafficking inhibition and acute hERG block

In this thesis, in Chapter 2, I have provided a comprehensive review of the pharmacological agents that have been reported to cause various cardiac complications in humans. A number of drugs are known to cause hERG channel deficiency not only by acute channel block, but also by the inhibition of its trafficking. At the same time, for some drugs, the effect on trafficking was shown to be more potent than the acute channel block.

In Chapter 7, I introduced the modelling of arsenic trioxide, a drug used in leukaemia treatment and with known common QT prolongation [380]. However, other agents can be explored in the current setting in future studies. With the coupling of acute hERG block (simple pore-block or more complex

dynamic simulations) with hERG trafficking modelling, those two modes of drug action can be explored in more detail and compared.

8.2.4 Simulation of trafficking-deficient mutations of hERG

Genetic mutations are among the factors that can cause a deficiency of cardiac ion channels. Importantly, hERG mutations are one of the major monogenic causes of life-threatening conditions, such as type-2 Long QT syndrome (LQTS2). This is characterised by profoundly delayed repolarisation of cardiomyocytes, therefore posing an increased risk of fatal arrhythmic events. Remarkably, most of such mutations are known to interfere with the trafficking of the channel [365]. The vast majority of LQTS2 mutations cause the inhibition of hERG protein synthesis, where a defective copy of the transcript has a stop codon, which terminates the protein synthesis, or a splice site mutation causing missplicing and the generation of incomplete proteins or RNA decay [365]. Other trafficking mutations include a reduction in folding efficiency or increased degradation [365–367].

Notably, it has been shown through in vitro experiments that multiple trafficking mutations can be reversed, either with temperature or with pharmacological agents [381]. A large-scale study of 291 LQTS2 KCNH2 mutations showed that incubation of HEK cells expressing mutated channels at 27°C or/and with the hERG blocker and pharmacological chaperone E4031 could lead to an improvement in the levels of mature hERG [382]. However, those improvements are dependent on the location of the mutation, with some mutations being completely uncorrectable and some mutations being only partially correctable. Therefore, hERG trafficking mutations can be separated into five classes: 1) wild-type-like, 2) correctable by incubation at 27°C, 3) correctable by E4031 alone, 4) correctable by both 27°C and E4031, and 5) not correctable by 27°C or E4031. Later, Kanner and colleagues showed that the increased levels of mature hERG caused by incubation at low-temperature do not always result in an increase in hERG surface expression [367]. Interestingly, Mehta and colleagues showed that incubation of hiPSCs

with a hERG trafficking mutation via the protease inhibitor (N-[N-(N-acetyl-L-leucyl)-L-leucyl]-L-norleucine) for 48 hours can increase the levels of mature hERG by 9 times, although it will still remain 2.5 times lower compared to the wild type. This resulted in a shortening of the action potential duration at 90% repolarisation from 1130 to 700 ms, compared to 600 ms for control [368].

The proposed hERG trafficking model, coupled with action potential models of cardiac electrophysiology, holds great potential to investigate the underlying mechanisms of the above described electrophysiological abnormalities of trafficking LQTS2 mutations. This includes investigations on whether a single mechanism or the combination of an inhibited maturation, accelerated degradation and reduced surface stability of the mutant channels, result in the prolonged action potential of these mutant cells.

8.2.5 Trafficking modelling of other channels

As listed in Chapter 2, hERG trafficking inhibition is not the only reported off-target effect of drugs, which is the case of celastrol, which is reported to inhibit trafficking of Kir2.1 and hERG [52].

In addition to drug-induced channel trafficking inhibition and mutations that cause a deficiency of channels on the membrane, the regulation of ion channel densities on the myocyte membranes has been proposed as a specific mechanism of tissue electrical remodelling in conditions such as heart failure and atrial fibrillation [383], or obesity [384]. One example is the decreased protein expression of Kv1.5, Kv4.3 and Kir3.1 channels in cardiac tissue in patients with atrial fibrillation [385]. Regulation of ion channel trafficking has also begun to gain attention as a new potential therapeutic target for antiarrhythmic treatment [386,387].

The proposed model of hERG trafficking described in Chapters 6 and 7 could be extended and adapted to account for trafficking of other channels for future studies.

8.2.6 Tissue and organ effect of trafficking inhibition

The studies in this thesis were focused on single-cell and populations of cells modelling. At the same time, pro-arrhythmic electrophysiological abnormalities can manifest differently on a tissue and organ level. In addition, the heart is a complex organ, and observed abnormalities are only markers of arrhythmias as real arrhythmic events occur on the organ level. An extension of this thesis could explore the effects of hERG trafficking inhibition by coupling the outputs of trafficking models with mechanistic tissue and organ simulations. Similarly to the work conducted in Chapter 7, at the individual cell level, multi-scale simulations can be performed to study the effects of hERG trafficking inhibitors in an *in silico* clinical trial setting at the whole-organ and body levels, to reconstruct ECG recordings and calculate more clinically relevant arrhythmic-risk biomarkers.

8.3 Conclusions

Cardiotoxicity is one of the most common and life-threatening drug-induced adverse effects. Given its importance in precipitating potentially lethal arrhythmias, current preclinical testing is largely focused on the detection of unforeseen electrophysiological drug effects. However, a substantial number of drugs also affect cardiac function at many other levels. In order to address such challenge, I have developed a computational machine learning model that goes beyond the prediction of single drug-induced adverse effects. To do this, I firstly collected a large dataset of 1131 drugs with their transcriptional profiles, structure and safety information. Using these data, I proposed a multi-label model that takes the structure and cellular gene expression profiles of drugs and predicts up to six cardiotoxicities simultaneously. I extensively validated this approach on drugs and cell lines previously unseen by the algorithm. I then moved to mechanistic modelling and focused on a specific and highly relevant mechanism of drug-induced cardiotoxicity: inhibition of hERG trafficking. I formulated a first-in-kind model of hERG trafficking and

used it to explore the effects of hERG trafficking inhibition on action potential properties. Employing a population of models approach, I identified and quantified the action of arsenic trioxide, a pharmacological agent known to inhibit trafficking, both at protein and electrophysiological levels. In conclusion, this thesis presents two novel and complementary approaches for drug safety analysis, and shows that computational methods have the potential to complement existing drug discovery methods.

9 Bibliography

1. The top 10 causes of death [Internet]. [cited 2020 Jan 11]. Available from: <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>
2. MacRae CA, Roden DM, Loscalzo J. The Future of Cardiovascular Therapeutics. *Circulation* [Internet]. 2016 [cited 2020 Jan 11]; 133: 2610–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27324356>
3. Fordyce CB, Roe MT, Ahmad T, Libby P, Borer JS, Hiatt WR, Bristow MR, Packer M, Wasserman SM, Braunstein N, Pitt B, DeMets DL, Cooper-Arnold K, et al. Cardiovascular Drug Development. *J Am Coll Cardiol* [Internet]. 2015 [cited 2020 Jan 11]; 65: 1567–82. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0735109715008232>
4. Mullard A. Parsing clinical success rates. *Nat Rev Drug Discov* [Internet]. 2016 [cited 2020 Jan 11]; 15: 447–447. Available from: <http://www.nature.com/articles/nrd.2016.136>
5. Onakpoya IJ, Heneghan CJ, Aronson JK. Post-marketing withdrawal of 462 medicinal products because of adverse drug reactions: a systematic review of the world literature. *BMC Med* [Internet]. England; 2016; 14: 10. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26843061>
6. Mak IW, Evaniew N, Ghert M. Lost in translation: animal models and clinical trials in cancer treatment. *Am J Transl Res* [Internet]. 2014 [cited 2019 Oct 11]; 6: 114–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24489990>
7. Cesarovic N, Lipiski M, Falk V, Emmert MY. Animals in cardiovascular research. *Eur Heart J* [Internet]. 2020 [cited 2020 Jan 11]; 41: 200–3. Available from: <https://academic.oup.com/eurheartj/article/41/2/200/5697525>
8. Qian T, Zhu S, Hoshida Y. Use of big data in drug development for precision medicine: an update. *Expert Review of Precision Medicine and Drug Development*. Taylor and Francis Ltd.; 2019. p. 189–200.
9. Lo AW, Siah KW, Wong CH. Machine Learning with Statistical Imputation for Predicting Drug Approval. *Harvard Data Sci Rev*

- [Internet]. 2019 [cited 2020 Jan 11]; . Available from: <https://hdsr.mitpress.mit.edu/pub/ct67j043>
10. Lipinski CF, Maltarollo VG, Oliveira PR, da Silva ABF, Honorio KM. Advances and Perspectives in Applying Deep Learning for Drug Design and Discovery. *Front Robot AI* [Internet]. 2019 [cited 2020 Jan 11]; 6. Available from: <https://www.frontiersin.org/article/10.3389/frobt.2019.00108/full>
 11. Vamathevan J, Clark D, Czodrowski P, Dunham I, Ferran E, Lee G, Li B, Madabhushi A, Shah P, Spitzer M, Zhao S. Applications of machine learning in drug discovery and development. *Nature Reviews Drug Discovery*. Nature Publishing Group; 2019. p. 463–77.
 12. Davies MR, Wang K, Mirams GR, Caruso A, Noble D, Walz A, Lavé T, Schuler F, Singer T, Polonchuk L. Recent developments in using mechanistic cardiac modelling for drug safety evaluation. *Drug Discovery Today*. Elsevier Ltd; 2016. p. 924–38.
 13. Sloan JA, Goldberg RM, Sargent DJ, Vargas-Chanes D, Nair S, Cha SS, Novotny PJ, Poon MA, O'Connell MJ, Loprinzi CL. Women experience greater toxicity with fluorouracil-based chemotherapy for colorectal cancer. *J Clin Oncol*. United States; 2002; 20: 1491–8.
 14. Volkova M, Russell R. Anthracycline Cardiotoxicity: Prevalence, Pathogenesis and Treatment. *Curr Cardiol Rev*. Bentham Science Publishers Ltd.; 2012; 7: 214–20.
 15. Garcia-Pavia P, Kim Y, Restrepo-Cordoba MA, Lunde IG, Wakimoto H, Smith AM, Toepfer CN, Getz K, Gorham J, Patel P, Ito K, Willcox JA, Arany Z, et al. Genetic Variants Associated With Cancer Therapy–Induced Cardiomyopathy. *Circulation*. Ovid Technologies (Wolters Kluwer Health); 2019; 140: 31–41.
 16. Ford GA, Wood SM, Daly AK. CYP2D6 and CYP2C19 genotypes of patients with terodiline cardiotoxicity identified through the yellow card system. *Br J Clin Pharmacol* [Internet]. 2000 [cited 2019 Sep 27]; 50: 77–80. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/10886124>
 17. Sajjad M, Fradley M, Sun W, Kim J, Zhao X, Pal T, Ismail-Khan R. An Exploratory Study to Determine Whether BRCA1 and BRCA2 Mutation Carriers Have Higher Risk of Cardiac Toxicity. Nock NL, editor. *Genes*

- (Basel) [Internet]. MDPI; 2017; 8: 59. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC5333048/>
18. Zordoky BN, Radin MJ, Heller L, Tobias A, Matise I, Apple FS, McCune SA, Sharkey LC. The interplay between genetic background and sexual dimorphism of doxorubicin-induced cardiotoxicity. *Cardio-Oncology*. Springer Nature; 2016; 2.
 19. Yap YG, Camm AJ. Drug induced QT prolongation and torsades de pointes. *Heart* [Internet]. 2003 [cited 2019 Sep 28]; 89: 1363–72. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/14594906>
 20. Friedrichs GS, Patmore L, Bass A. Non-clinical evaluation of ventricular repolarization (ICH S7B): Results of an interim survey of international pharmaceutical companies. *J Pharmacol Toxicol Methods* [Internet]. 2005; 52: 6–11. Available from: <http://www.sciencedirect.com/science/article/pii/S1056871905000699>
 21. Babcock JJ, Li M. hERG channel function: beyond long QT. *Acta Pharmacol Sin*. United States; 2013; 34: 329–35.
 22. Arcangeli A, Becchetti A. hERG Channels: From Antitargets to Novel Targets for Cancer Therapy. *Clinical cancer research : an official journal of the American Association for Cancer Research*. United States; 2017. p. 3–5.
 23. Brennan T, Fink M, Rodriguez B. Multiscale modelling of drug-induced effects on cardiac electrophysiological activity. *Eur J Pharm Sci* [Internet]. 2009 [cited 2019 Sep 27]; 36: 62–77. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19061955>
 24. Hoffmann P, Warner B. Are hERG channel inhibition and QT interval prolongation all there is in drug-induced torsadogenesis? A review of emerging trends. *J Pharmacol Toxicol Methods* [Internet]. 2006; 53: 87–105. Available from: <http://www.sciencedirect.com/science/article/pii/S1056871905000936>
 25. Mirams GR, Cui Y, Sher A, Fink M, Cooper J, Heath BM, McMahon NC, Gavaghan DJ, Noble D. Simulation of multiple ion channel block provides improved early prediction of compounds' clinical torsadogenic risk. *Cardiovasc Res* [Internet]. Oxford University Press; 2011; 91: 53–61. Available from:

- <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3112019/>
26. Kramer J, Obejero-Paz CA, Myatt G, Kuryshev YA, Bruening-Wright A, Verducci JS, Brown AM. MICE models: Superior to the HERG model in predicting torsade de pointes. *Sci Rep*. 2013; 3.
 27. Britton OJ, Bueno-Orovio A, Virág L, Varró A, Rodriguez B. The electrogenic Na⁺/K⁺ pump is a key determinant of repolarization abnormality susceptibility in human ventricular cardiomyocytes: A population-based simulation study. *Front Physiol*. Frontiers Media S.A.; 2017; 8.
 28. Li X, Zhang R, Zhao B, Lossin C, Cao Z. Cardiotoxicity screening: a review of rapid-throughput in vitro approaches. *Arch Toxicol* [Internet]. 2016 [cited 2019 Sep 27]; 90: 1803–16. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26676948>
 29. Lu HR, Vlamincx E, Hermans AN, Rohrbacher J, Van Ammel K, Towart R, Pugsley M, Gallacher DJ. Predicting drug-induced changes in QT interval and arrhythmias: QT-shortening drugs point to gaps in the ICHS7B Guidelines. *Br J Pharmacol*. 2008; 154: 1427–38.
 30. Turan B, Dhalla SN, editors. *Diabetic Cardiomyopathy*. Springer Science & Business Media; 2014.
 31. Nogawa H, Kawai T. hERG trafficking inhibition in drug-induced lethal cardiac arrhythmia. *Eur J Pharmacol*. Netherlands; 2014; 741: 336–9.
 32. Cubeddu LX. Drug-induced Inhibition and Trafficking Disruption of ion Channels: Pathogenesis of QT Abnormalities and Drug-induced Fatal Arrhythmias. *Curr Cardiol Rev*. United Arab Emirates; 2016; 12: 141–54.
 33. Ficker E, Kuryshev YA, Dennis AT, Obejero-Paz C, Wang L, Hawryluk P, Wible BA, Brown AM. Mechanisms of arsenic-induced prolongation of cardiac repolarization. *Mol Pharmacol*. United States; 2004; 66: 33–44.
 34. Zhang X, Yu D, Geng H, Li F, Lv L, Zhao L, Yan C, Li B. Dual effects of arsenic trioxide on tumor cells and the potential underlying mechanisms. *Oncol Lett*. Greece; 2018; 16: 3812–20.
 35. Zhang K, Zhi D, Huang T, Gong Y, Yan M, Liu C, Wei T, Dong Z, Li B, Yang B. Berberine induces hERG channel deficiency through trafficking

- inhibition. *Cell Physiol Biochem*. 2014; 34: 691–702.
36. Yu D, Lv L, Fang L, Zhang B, Wang J, Zhan G, Zhao L, Zhao X, Li B. Inhibitory effects and mechanism of dihydroberberine on hERG channels expressed in HEK293 cells. Xu S-Z, editor. *PLoS One* [Internet]. San Francisco, CA USA: Public Library of Science; 2017; 12: e0181823. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC5538702/>
 37. Weng C-Y, Chiou S-Y, Wang L, Kou M-C, Wang Y-J, Wu M-J. Arsenic trioxide induces unfolded protein response in vascular endothelial cells. *Arch Toxicol. Germany*; 2014; 88: 213–26.
 38. Dennis AT, Wang L, Wan H, Nassal D, Deschenes I, Ficker E. Molecular determinants of pentamidine-induced hERG trafficking inhibition. *Mol Pharmacol. United States*; 2012; 81: 198–209.
 39. Wang L, Wible BA, Wan X, Ficker E. Cardiac glycosides as novel inhibitors of human ether-a-go-go-related gene channel trafficking. *J Pharmacol Exp Ther. United States*; 2007; 320: 525–34.
 40. Wang L, Dennis AT, Trieu P, Charron F, Ethier N, Hebert TE, Wan X, Ficker E. Intracellular potassium stabilizes human ether-a-go-go-related gene channels for export from endoplasmic reticulum. *Mol Pharmacol. United States*; 2009; 75: 927–37.
 41. Apaja PM, Foo B, Okiyoneda T, Valinsky WC, Barriere H, Atanasiu R, Ficker E, Lukacs GL, Shrier A. Ubiquitination-dependent quality control of hERG K⁺ channel with acquired and inherited conformational defect at the plasma membrane. *Mol Biol Cell. United States*; 2013; 24: 3787–804.
 42. Hubbi ME, Hu H, Kshitiz, Ahmed I, Levchenko A, Semenza GL. Chaperone-mediated autophagy targets hypoxia-inducible factor-1alpha (HIF-1alpha) for lysosomal degradation. *J Biol Chem. United States*; 2013; 288: 10703–14.
 43. Guo J, Massaeli H, Li W, Xu J, Luo T, Shaw J, Kirshenbaum LA, Zhang S. Identification of IKr and its trafficking disruption induced by probucol in cultured neonatal rat cardiomyocytes. *J Pharmacol Exp Ther. United States*; 2007; 321: 911–20.
 44. Shi Y-Q, Yan C-C, Zhang X, Yan M, Liu L-R, Geng H-Z, Lv L, Li B-X.

- Mechanisms underlying probucol-induced hERG-channel deficiency. *Drug Des Devel Ther.* New Zealand; 2015; 9: 3695–704.
45. Spence S, Deurinck M, Ju H, Traebert M, McLean L, Marlowe J, Emotte C, Tritto E, Tseng M, Shultz M, Friedrichs GS. Histone Deacetylase Inhibitors Prolong Cardiac Repolarization through Transcriptional Mechanisms. *Toxicol Sci.* United States; 2016; 153: 39–54.
 46. Takemasa H, Nagatomo T, Abe H, Kawakami K, Igarashi T, Tsurugi T, Kabashima N, Tamura M, Okazaki M, Delisle BP, January CT, Otsuji Y. Coexistence of hERG current block and disruption of protein trafficking in ketoconazole-induced long QT syndrome. *Br J Pharmacol.* England; 2008; 153: 439–47.
 47. Han S, Sun X, Zhang Z, Zhang L. The protease inhibitor atazanavir blocks hERG K(+) channels expressed in HEK293 cells and obstructs hERG protein transport to cell membrane. *Acta Pharmacol Sin.* United States; 2015; 36: 454–62.
 48. Rajamani S, Eckhardt LL, Valdivia CR, Klemens CA, Gillman BM, Anderson CL, Holzem KM, Delisle BP, Anson BD, Makielski JC, January CT. Drug-induced long QT syndrome: hERG K(+) channel block and disruption of protein trafficking by fluoxetine and norfluoxetine. *Br J Pharmacol* [Internet]. Nature Publishing Group; 2006; 149: 481–9. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2014667/>
 49. Rajamani S, Eckhardt LL, Valdivia CR, Klemens CA, Gillman BM, Anderson CL, Holzem KM, Delisle BP, Anson BD, Makielski JC, January CT. Drug-induced long QT syndrome: hERG K+ channel block and disruption of protein trafficking by fluoxetine and norfluoxetine. *Br J Pharmacol* [Internet]. 2006 [cited 2019 Sep 27]; 149: 481–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16967046>
 50. Ofek K, Schoknecht K, Melamed-Book N, Heinemann U, Friedman A, Soreq H. Fluoxetine induces vasodilatation of cerebral arterioles by co-modulating NO/muscarinic signalling. *J Cell Mol Med.* England; 2012; 16: 2736–44.
 51. Dennis AT, Nassal D, Deschenes I, Thomas D, Ficker E. Antidepressant-induced ubiquitination and degradation of the cardiac

- potassium channel hERG. *J Biol Chem. United States*; 2011; 286: 34413–25.
52. Sun H, Liu X, Xiong Q, Shikano S, Li M. Chronic inhibition of cardiac Kir2.1 and HERG potassium channels by celastrol with dual effects on both ion conductivity and protein trafficking. *J Biol Chem. United States*; 2006; 281: 5877–84.
 53. Dewenter M, von der Lieth A, Katus HA, Backs J. Calcium Signaling and Transcriptional Regulation in Cardiomyocytes. *Circ Res [Internet]*. 2017 [cited 2019 Sep 27]; 121: 1000–20. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28963192>
 54. Mustroph J, Neef S, Maier LS. CaMKII as a target for arrhythmia suppression. *Pharmacol Ther. England*; 2017; 176: 22–31.
 55. Hood AR, Ai X, Pogwizd SM. Regulation of cardiac gap junctions by protein phosphatases. *Journal of Molecular and Cellular Cardiology. Academic Press*; 2017. p. 52–7.
 56. Yu L-Y, Sun L-N, Zhang X-H, Li Y-Q, Yu L, Yuan Z-Q-Y, Meng L, Zhang H-W, Wang Y-Q. A Review of the Novel Application and Potential Adverse Effects of Proton Pump Inhibitors. *Adv Ther [Internet]*. 2017 [cited 2019 Sep 27]; 34: 1070–86. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28429247>
 57. Mooney L, Skinner M, Coker SJ, Currie S. Effects of acute and chronic sunitinib treatment on cardiac function and calcium/calmodulin-dependent protein kinase II. *Br J Pharmacol. John Wiley and Sons Inc.*; 2015; 172: 4342–54.
 58. Rosano GMC, Fini M, Caminiti G, Barbaro G. Cardiac metabolism in myocardial ischemia. *Curr Pharm Des [Internet]*. 2008 [cited 2019 Sep 27]; 14: 2551–62. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18991672>
 59. Williams LJ, Nye BG, Wende AR. Diabetes-Related Cardiac Dysfunction. *Endocrinol Metab (Seoul, Korea) [Internet]*. 2017 [cited 2019 Sep 27]; 32: 171–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28685508>
 60. Ott M, Gogvadze V, Orrenius S, Zhivotovsky B. Mitochondria, oxidative stress and cell death. *Apoptosis [Internet]*. 2007 [cited 2019 Sep 27];

- 12: 913–22. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17453160>
61. Youle RJ, van der Bliek AM. Mitochondrial fission, fusion, and stress. *Science* [Internet]. 2012 [cited 2019 Sep 27]; 337: 1062–5. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22936770>
 62. Eisner V, Cupo RR, Gao E, Csordás G, Slovinsky WS, Paillard M, Cheng L, Ibeti J, Chen SRW, Chuprun JK, Hoek JB, Koch WJ, Hajnóczky G. Mitochondrial fusion dynamics is robust in the heart and depends on calcium oscillations and contractile activity. *Proc Natl Acad Sci U S A* [Internet]. 2017 [cited 2019 Sep 27]; 114: E859–68. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28096338>
 63. Givvimani S, Pushpakumar SB, Metreveli N, Veeranki S, Kundu S, Tyagi SC. Role of mitochondrial fission and fusion in cardiomyocyte contractility. *Int J Cardiol* [Internet]. 2015 [cited 2019 Sep 27]; 187: 325–33. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25841124>
 64. Varga Z V, Ferdinandy P, Liaudet L, Pacher P. Drug-induced mitochondrial dysfunction and cardiotoxicity. *Am J Physiol Heart Circ Physiol* [Internet]. 2015 [cited 2019 Sep 27]; 309: H1453–67. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26386112>
 65. McGowan J V, Chung R, Maulik A, Piotrowska I, Walker JM, Yellon DM. Anthracycline Chemotherapy and Cardiotoxicity. *Cardiovasc drugs Ther* [Internet]. 2017 [cited 2019 Sep 27]; 31: 63–75. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28185035>
 66. Lewis W, Day BJ, Copeland WC. Mitochondrial toxicity of NRTI antiviral drugs: an integrated cellular perspective. *Nat Rev Drug Discov* [Internet]. 2003 [cited 2019 Sep 27]; 2: 812–22. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/14526384>
 67. Nomura R, Sato T, Sato Y, Medin JA, Kushimoto S, Yanagisawa T. Azidothymidine-triphosphate impairs mitochondrial dynamics by disrupting the quality control system. *Redox Biol* [Internet]. 2017 [cited 2019 Sep 27]; 13: 407–17. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28683400>
 68. Hiller N, Mirtschink P, Merkel C, Knels L, Oertel R, Christ T, Deussen A, Koch T, Stehr SN. Myocardial accumulation of bupivacaine and

- ropivacaine is associated with reversible effects on mitochondria and reduced myocardial function. *Anesth Analg* [Internet]. 2013 [cited 2019 Sep 27]; 116: 83–92. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23223114>
69. Tsuchiya H, Mizogami M. Interaction of local anesthetics with biomembranes consisting of phospholipids and cholesterol: Mechanistic and clinical implications for anesthetic and cardiotoxic effects. *Anesthesiology Research and Practice*. 2013.
 70. Gentile C, Chuang CY, Drake CJ, Davies MJ. VEGF/eNOS-mediated Cardiomyocyte Proliferation: a Novel Mechanism to Regulate Cardiovascular Regeneration in Mice and Humans. *FASEB J* [Internet]. Federation of American Societies for Experimental Biology; 2016; 30: 1b568–1b568. Available from: https://www.fasebj.org/doi/abs/10.1096/fasebj.30.1_supplement.1b568
 71. Taimeh Z, Loughran J, Birks EJ, Bolli R. Vascular endothelial growth factor in heart failure. *Nat Rev Cardiol* [Internet]. 2013 [cited 2019 Sep 27]; 10: 519–30. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23856679>
 72. Zhao T, Zhao W, Meng W, Liu C, Chen Y, Gerling IC, Weber KT, Bhattacharya SK, Kumar R, Sun Y. VEGF-C/VEGFR-3 pathway promotes myocyte hypertrophy and survival in the infarcted myocardium. *Am J Transl Res* [Internet]. 2015 [cited 2019 Sep 27]; 7: 697–709. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26064438>
 73. Quesada AR, Medina MÁ. Angiogenesis in cancer: Therapeutic targets and angiogenesis inhibitors. *Angiogenesis & Therapeutic Targets In Cancer*. Bentham Science Publishers Ltd.; 2010. p. 45–67.
 74. Zi C, Ai D. Cardiotoxicity associated with targeted cancer therapies. *Mol Clin Oncol*. Spandidos Publications; 2016; 4: 675–81.
 75. De Luisi A, Mangialardi G, Ria R, Acuto G, Ribatti D, Vacca A. Anti-angiogenic activity of carebistine: a plausible mechanism affecting airway remodelling. *Eur Respir J* [Internet]. 2009 [cited 2019 Sep 27]; 34: 958–66. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19357149>

76. Lu Q, Wang C, Pan R, Gao X, Wei Z, Xia Y, Dai Y. Histamine synergistically promotes bFGF-induced angiogenesis by enhancing VEGF production via H1 receptor. *J Cell Biochem* [Internet]. 2013 [cited 2019 Sep 27]; 114: 1009–19. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23225320>
77. Negro A, Brar BK, Lee K-F. Essential roles of Her2/erbB2 in cardiac development and function. *Recent Prog Horm Res* [Internet]. 2004 [cited 2019 Sep 27]; 59: 1–12. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/14749494>
78. Makki N, Thiel KW, Miller FJ. The epidermal growth factor receptor and its ligands in cardiovascular disease. *Int J Mol Sci* [Internet]. 2013 [cited 2019 Sep 27]; 14: 20597–613. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24132149>
79. Harari PM. Epidermal growth factor receptor inhibition strategies in oncology. *Endocr Relat Cancer* [Internet]. 2004 [cited 2019 Sep 27]; 11: 689–708. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15613446>
80. Pondé NF, Lambertini M, de Azambuja E. Twenty years of anti-HER2 therapy-associated cardiotoxicity. *ESMO open* [Internet]. 2016 [cited 2019 Sep 27]; 1: e000073. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27843627>
81. Bloomekatz J, Singh R, Prall OW, Dunn AC, Vaughan M, Loo C-S, Harvey RP, Yelon D. Platelet-derived growth factor (PDGF) signaling directs cardiomyocyte movement toward the midline during heart tube assembly. *Elife*. England; 2017; 6.
82. Chintalgattu V, Ai D, Langley RR, Zhang J, Bankson JA, Shih TL, Reddy AK, Coombes KR, Daher IN, Pati S, Patel SS, Pocius JS, Taffet GE, et al. Cardiomyocyte PDGFR-beta signaling is an essential component of the mouse cardiac response to load-induced stress. *J Clin Invest* [Internet]. 2010 [cited 2019 Sep 27]; 120: 472–84. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20071776>
83. Gallini R, Lindblom P, Bondjers C, Betsholtz C, Andrae J. PDGF-A and PDGF-B induces cardiac fibrosis in transgenic mice. *Exp Cell Res* [Internet]. 2016 [cited 2019 Sep 27]; 349: 282–90. Available from:

- <http://www.ncbi.nlm.nih.gov/pubmed/27816607>
84. Toyama EQ, Herzig S, Courchet J, Lewis TL, Losón OC, Hellberg K, Young NP, Chen H, Polleux F, Chan DC, Shaw RJ. Metabolism: AMP-activated protein kinase mediates mitochondrial fission in response to energy stress. *Science* (80-). American Association for the Advancement of Science; 2016; 351: 275–81.
 85. Curigliano G, Cardinale D, Dent S, Criscitiello C, Aseyev O, Lenihan D, Cipolla CM. Cardiotoxicity of anticancer treatments: Epidemiology, detection, and management. *CA Cancer J Clin* [Internet]. 2016 [cited 2019 Sep 27]; 66: 309–25. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26919165>
 86. Ewer MS, Ewer SM. Cardiotoxicity of anticancer treatments. *Nat Rev Cardiol* [Internet]. 2015 [cited 2019 Sep 27]; 12: 547–58. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25962976>
 87. Mitry MA, Edwards JG. Doxorubicin induced heart failure: Phenotype and molecular mechanisms. *IJC Heart and Vasculature*. Elsevier Ireland Ltd; 2016. p. 17–24.
 88. Mantawy EM, El-Bakly WM, Esmat A, Badr AM, El-Demerdash E. Chrysin alleviates acute doxorubicin cardiotoxicity in rats via suppression of oxidative stress, inflammation and apoptosis. *Eur J Pharmacol* [Internet]. 2014 [cited 2019 Sep 27]; 728: 107–18. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24509133>
 89. Geisberg CA, Sawyer DB. Mechanisms of anthracycline cardiotoxicity and strategies to decrease cardiac damage. *Curr Hypertens Rep* [Internet]. 2010 [cited 2019 Sep 27]; 12: 404–10. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20842465>
 90. Zhao L, Zhang B. Doxorubicin induces cardiotoxicity through upregulation of death receptors mediated apoptosis in cardiomyocytes. *Sci Rep*. England; 2017; 7: 44735.
 91. Polk A, Vaage-Nilsen M, Vistisen K, Nielsen DL. Cardiotoxicity in cancer patients treated with 5-fluorouracil or capecitabine: a systematic review of incidence, manifestations and predisposing factors. *Cancer Treat Rev* [Internet]. 2013 [cited 2019 Sep 27]; 39: 974–84. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23582737>

92. Van Cutsem E, Hoff PM, Blum JL, Abt M, Osterwalder B. Incidence of cardiotoxicity with the oral fluoropyrimidine capecitabine is typical of that reported with 5-fluorouracil. *Ann Oncol Off J Eur Soc Med Oncol* [Internet]. 2002 [cited 2019 Sep 27]; 13: 484–5. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/11996484>
93. Iorio-Morin C, Germain P, Roy S, Génier S, Labrecque P, Parent J-L. Thromboxane A₂ modulates cisplatin-induced apoptosis through a Siva1-dependent mechanism. *Cell Death Differ* [Internet]. 2012 [cited 2019 Sep 27]; 19: 1347–57. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22343716>
94. Togna GI, Togna AR, Franconi M, Caprino L. Cisplatin triggers platelet activation. *Thromb Res*. Elsevier Ltd; 2000; 99: 503–9.
95. McQuade RM, Stojanovska V, Abalo R, Bornstein JC, Nurgali K. Chemotherapy-Induced Constipation and Diarrhea: Pathophysiology, Current and Emerging Treatments. *Front Pharmacol* [Internet]. 2016 [cited 2019 Sep 27]; 7: 414. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27857691>
96. Kucharz J, Michalowska-Kaczmarczyk A, Zygulska AL, Wojtak J, Pawlik W, Herman RM, Krzemieniecki K. Bradycardia as a rare symptom of cisplatin cardiotoxicity: A case report. *Oncol Lett* [Internet]. 2016 [cited 2019 Sep 27]; 11: 2297–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26998165>
97. Conrad ME. Cytarabine and cardiac failure. *Am J Hematol* [Internet]. 1992 [cited 2019 Sep 27]; 41: 143–4. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/1415181>
98. Wayangankar SA, Patel BC, Parekh HD, Holter JL, Lazzara R. High-dose cytosine arabinoside-induced symptomatic bradycardia. *J Cardiovasc Med (Hagerstown)* [Internet]. 2015 [cited 2019 Sep 27]; 16 Suppl 1: S38–41. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21178637>
99. Bergeron A, Raffy O, Vannetzel JM. Myocardial ischemia and infarction associated with vinorelbine. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. United States; 1995. p. 531–2.

100. Suter TM, Ewer MS. Cancer drugs and the heart: importance and management. *Eur Heart J* [Internet]. 2013 [cited 2019 Sep 27]; 34: 1102–11. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22789916>
101. Mikaelian I, Buness A, de Vera-Mudry M-C, Kanwal C, Coluccio D, Rasmussen E, Char HW, Carvajal V, Hilton H, Funk J, Hoflack J-C, Fielden M, Herting F, et al. Primary endothelial damage is the mechanism of cardiotoxicity of tubulin-binding drugs. *Toxicol Sci* [Internet]. 2010 [cited 2019 Sep 27]; 117: 144–51. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20624997>
102. Schimmel KJM, Richel DJ, van den Brink RBA, Guchelaar H-J. Cardiotoxicity of cytotoxic drugs. *Cancer Treat Rev* [Internet]. 2004 [cited 2019 Sep 27]; 30: 181–91. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15023436>
103. Arora A, Scholar EM. Role of tyrosine kinase inhibitors in cancer therapy. *J Pharmacol Exp Ther*. United States; 2005; 315: 971–9.
104. Chen MH, Kerkelä R, Force T. Mechanisms of cardiac dysfunction associated with tyrosine kinase inhibitor cancer therapeutics. *Circulation* [Internet]. 2008 [cited 2019 Sep 27]; 118: 84–95. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18591451>
105. Lenihan DJ, Kowey PR. Overview and Management of Cardiac Adverse Events Associated With Tyrosine Kinase Inhibitors. *Oncologist*. Alphamed Press; 2013; 18: 900–8.
106. Rea D. Management of adverse events associated with tyrosine kinase inhibitors in chronic myeloid leukemia. *Ann Hematol* [Internet]. 2015 [cited 2019 Sep 27]; 94 Suppl 2: S149-58. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25814081>
107. Force T, Kolaja KL. Cardiotoxicity of kinase inhibitors: The prediction and translation of preclinical models to clinical outcomes. *Nature Reviews Drug Discovery*. 2011. p. 111–26.
108. Doherty KR, Talbert DR, Trusk PB, Moran DM, Shell SA, Bacus S. Structural and functional screening in human induced-pluripotent stem cell-derived cardiomyocytes accurately identifies cardiotoxicity of multiple drug types. *Toxicol Appl Pharmacol*. Academic Press Inc.;

- 2015; 285: 51–60.
109. Xu Z, Cang S, Yang T, Liu D. Cardiotoxicity of tyrosine kinase inhibitors in chronic myelogenous leukemia therapy. *Hematol Rep. PAGEPress Publications*; 2009; 1: 4.
 110. Moslehi JJ, Deininger M. Tyrosine kinase inhibitor-associated cardiovascular toxicity in chronic myeloid leukemia. *Journal of Clinical Oncology. American Society of Clinical Oncology*; 2015. p. 4210–8.
 111. Dias A, Claudino W, Sinha R, Perez CA, Jain D. Human epidermal growth factor antagonists and cardiotoxicity-A short review of the problem and preventative measures. *Crit Rev Oncol Hematol [Internet]*. 2016 [cited 2019 Sep 27]; 104: 42–51. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27338847>
 112. Zeglinski M, Ludke A, Jassal DS, Singal PK. Trastuzumab-induced cardiac dysfunction: A “dual-hit”. *Exp Clin Cardiol [Internet]*. 2011 [cited 2019 Sep 27]; 16: 70–4. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22065936>
 113. Economopoulou P, Kotsakis A, Kapisir I, Kentepozidis N. Cancer therapy and cardiovascular risk: focus on bevacizumab. *Cancer Manag Res [Internet]*. 2015 [cited 2019 Sep 27]; 7: 133–43. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26082660>
 114. Lancellotti P, Zamorano JL, Galderisi M. Anticancer treatments and cardiotoxicity: mechanisms, diagnostic and therapeutic interventions. *Academic Press is an imprint of Elsevier*; 2016.
 115. Gryder BE, Sodji QH, Oyelere AK. Targeted cancer therapy: giving histone deacetylase inhibitors all they need to succeed. *Future Med Chem [Internet]*. 2012 [cited 2019 Sep 27]; 4: 505–24. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22416777>
 116. Winnik S, Auwerx J, Sinclair DA, Matter CM. Protective effects of sirtuins in cardiovascular diseases: from bench to bedside. *Eur Heart J [Internet]*. 2015 [cited 2019 Sep 27]; 36: 3404–12. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26112889>
 117. Mohamed Saleem TS, Basha SD. Red wine: A drink to your heart. *Journal of Cardiovascular Disease Research. Medknow Publications*; 2010. p. 171–6.

118. Meseeha MG, Kolade VO, Attia MN. Partially reversible bortezomib-induced cardiotoxicity: an unusual cause of acute cardiomyopathy. *J Community Hosp Intern Med Perspect*. Informa UK Limited; 2015; 5: 28982.
119. Xiao Y, Yin J, Wei J, Shang Z. Incidence and risk of cardiotoxicity associated with bortezomib in the treatment of cancer: A systematic review and meta-analysis. *PLoS ONE*. Public Library of Science; 2014.
120. Nowis D, Mączewski M, Mackiewicz U, Kujawa M, Ratajska A, Wieckowski MR, Wilczyński GM, Malinowska M, Bil J, Salwa P, Bugajski M, Wójcik C, Siński M, et al. Cardiotoxicity of the anticancer therapeutic agent bortezomib. *Am J Pathol*. American Society for Investigative Pathology Inc.; 2010; 176: 2658–68.
121. Johnson DB, Balko JM, Compton ML, Chalkias S, Gorham J, Xu Y, Hicks M, Puzanov I, Alexander MR, Bloomer TL, Becker JR, Slosky DA, Phillips EJ, et al. Fulminant myocarditis with combination immune checkpoint blockade. *N Engl J Med*. Massachusetts Medical Society; 2016; 375: 1749–55.
122. Varricchi G, Marone G, Mercurio V, Rosaria Galdiero M, Bonaduce D, G. Tocchetti C. Immune Checkpoint Inhibitors and Cardiac Toxicity: An Emerging Issue. *Curr Med Chem* [Internet]. Netherlands; 2017; 24. Available from: <http://www.ingentaconnect.com/content/10.2174/0929867324666170407125017>
123. Kourie HR, Klastersky J. Immune checkpoint inhibitors side effects and management. *Immunotherapy* [Internet]. 2016 [cited 2019 Sep 28]; 8: 799–807. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27349979>
124. Unnikrishnan D, Dutcher JP, Garl S, Varshneya N, Lucariello R, Wiernik PH. Cardiac monitoring of patients receiving arsenic trioxide therapy. *Br J Haematol*. 2004; 124: 610–7.
125. Ohnishi K, Yoshida H, Shigeno K, Nakamura S, Fujisawa S, Naito K, Shinjo K, Fujita Y, Matsui H, Takeshita A, Sugiyama S, Satoh H, Terada H, et al. Prolongation of the QT interval and ventricular tachycardia in patients treated with arsenic trioxide for acute promyelocytic leukemia.

- Ann Intern Med. American College of Physicians; 2000; 133: 881–5.
126. Ducas RA, Seftel MD, Ducas J, Seifer C. Monomorphic ventricular tachycardia caused by arsenic trioxide therapy for acute promyelocytic leukaemia. J R Coll Physicians Edinb [Internet]. 2011 [cited 2019 Sep 29]; 41: 117–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21677914>
 127. Yan M, Feng L, Shi Y, Wang J, Liu Y, Li F, Li B. Mechanism of As₂O₃-Induced Action Potential Prolongation and Using hiPS-CMs to Evaluate the Rescue Efficacy of Drugs With Different Rescue Mechanism. Toxicol Sci. United States; 2017; 158: 379–90.
 128. Christie LE, Picard J, Weinberg GL. Local anaesthetic systemic toxicity. BJA Educ [Internet]. 2015 [cited 2019 Sep 27]; 15: 136–42. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2058534917301543>
 129. Ruetsch YA, Böni T, Borgeat A. From cocaine to ropivacaine: the history of local anesthetic drugs. Curr Top Med Chem [Internet]. 2001 [cited 2019 Sep 27]; 1: 175–82. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/11895133>
 130. Schwartz BG, Rezkalla S, Kloner RA. Cardiovascular effects of cocaine. Circulation [Internet]. 2010 [cited 2019 Sep 27]; 122: 2558–69. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21156654>
 131. Goyal R, Shukla RN. Local anesthetic systemic toxicity (LAST) - should we not be concerned? Medical Journal Armed Forces India. Medical Journal Armed Forces India; 2012. p. 371–5.
 132. Graf BM. The cardiotoxicity of local anesthetics: the place of ropivacaine. Curr Top Med Chem [Internet]. 2001 [cited 2019 Sep 27]; 1: 207–14. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/11895138>
 133. Horáček M, Vymazal T. Lidocaine not so innocent: Cardiotoxicity after topical anaesthesia for bronchoscopy. Indian Journal of Anaesthesia. 2012. p. 95–6.
 134. Thanacoody HKR, Thomas SHL. Tricyclic antidepressant poisoning : cardiovascular toxicity. Toxicol Rev [Internet]. 2005 [cited 2019 Sep 28]; 24: 205–14. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16390222>

135. Pacher P, Kecskemeti V. Cardiovascular side effects of new antidepressants and antipsychotics: new drugs, old concerns? *Curr Pharm Des* [Internet]. 2004 [cited 2019 Sep 27]; 10: 2463–75. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15320756>
136. Sicouri S, Antzelevitch C. Sudden cardiac death secondary to antidepressant and antipsychotic drugs. *Expert Opinion on Drug Safety*. 2008. p. 181–94.
137. Khalifa M, Drolet B, Daleau P, Lefez C, Gilbert M, Plante S, O'Hara GE, Gleeton O, Hamelin BA, Turgeon J. Block of potassium currents in guinea pig ventricular myocytes and lengthening of cardiac repolarization in man by the histamine H1 receptor antagonist diphenhydramine. *J Pharmacol Exp Ther* [Internet]. 1999 [cited 2019 Sep 27]; 288: 858–65. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9918600>
138. Khalifa M, Daleau P, Turgeon J. Mechanism of sodium channel block by venlafaxine in guinea pig ventricular myocytes. *J Pharmacol Exp Ther* [Internet]. 1999 [cited 2019 Sep 27]; 291: 280–4. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/10490914>
139. Lenkey N, Karoly R, Lukacs P, Vizi ES, Sunesen M, Fodor L, Mike A. Classification of drugs based on properties of sodium channel inhibition: a comparative automated patch-clamp study. *PLoS One* [Internet]. 2010 [cited 2019 Sep 27]; 5: e15568. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21187965>
140. Foguet-Boreu Q, Fernandez San Martin MI, Flores Mateo G, Zabaleta Del Olmo E, Ayerbe García-Morzon L, Perez-Piñar López M, Martín-López LM, Montes Hidalgo J, Violán C. Cardiovascular risk assessment in patients with a severe mental illness: a systematic review and meta-analysis. *BMC Psychiatry* [Internet]. 2016 [cited 2019 Sep 27]; 16: 141. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27176477>
141. Hayward C. Psychiatric illness and cardiovascular disease risk. *Epidemiol Rev* [Internet]. 1995 [cited 2019 Sep 27]; 17: 129–38. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/8521931>
142. Larsen BA, Christenfeld NJS. Cardiovascular Disease and Psychiatric Comorbidity: The Potential Role of Perseverative Cognition. *Cardiovasc*

- Psychiatry Neurol. Hindawi Limited; 2009; 2009: 1–8.
143. Leucht S, Cipriani A, Spineli L, Mavridis D, Orey D, Richter F, Samara M, Barbui C, Engel RR, Geddes JR, Kissling W, Stapf MP, Lässig B, et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet* (London, England) [Internet]. 2013 [cited 2019 Sep 27]; 382: 951–62. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23810019>
 144. Hatton JL, Bhat PK, Gandhi S. Clozapine-induced myocarditis: recognizing a potentially fatal adverse reaction. *Texas Hear Inst J* [Internet]. 2015 [cited 2019 Sep 27]; 42: 155–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25873829>
 145. Layland JJ, Liew D, Prior DL. Clozapine-induced cardiotoxicity: a clinical update. *Med J Aust* [Internet]. 2009 [cited 2019 Sep 27]; 190: 190–2. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19220183>
 146. WHO (World Health Organization). Sertindole - approval application withdrawn. *WHO Pharmaceuticals Newsletter*. 1998; : 18.
 147. Lindström E, Farde L, Eberhard J, Haverkamp W. QTc interval prolongation and antipsychotic drug treatments: focus on sertindole. *Int J Neuropsychopharmacol* [Internet]. 2005 [cited 2019 Sep 27]; 8: 615–29. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15963244>
 148. Glassman AH. Schizophrenia, antipsychotic drugs, and cardiovascular disease. *J Clin Psychiatry*. United States; 2005; 66 Suppl 6: 5–10.
 149. Beach SR, Celano CM, Noseworthy PA, Januzzi JL, Huffman JC. QTc prolongation, torsades de pointes, and psychotropic medications. *Psychosomatics*. England; 2013; 54: 1–13.
 150. Yang T, Chun YW, Stroud DM, Mosley JD, Knollmann BC, Hong C, Roden DM. Screening for acute IKr block is insufficient to detect torsades de pointes liability: role of late sodium current. *Circulation* [Internet]. 2014 [cited 2019 Sep 27]; 130: 224–34. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24895457>
 151. Nasrallah HA. Atypical antipsychotic-induced metabolic side effects: insights from receptor-binding profiles. *Mol Psychiatry* [Internet]. 2008 [cited 2019 Sep 27]; 13: 27–35. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17848919>

152. Endoh M. Cardiac α 1-Adrenoceptors and Inotropy: Myofilament Ca^{2+} Sensitivity, Intracellular Ca^{2+} Mobilization, Signaling Pathway, and Pathophysiological Relevance. *Circulation Research*. Lippincott Williams and Wilkins; 2016. p. 587–90.
153. Shaunak S, Wilkins A, Pilling JB, Dick DJ. Pericardial, retroperitoneal, and pleural fibrosis induced by pergolide. *J Neurol Neurosurg Psychiatry*. BMJ Publishing Group; 1999; 66: 79–81.
154. Waller EA, Kaplan J. Pergolide-associated valvular heart disease. *Compr Ther* [Internet]. 2006 [cited 2019 Sep 27]; 32: 94–101. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16845152>
155. Antonini A, Poewe W. Fibrotic heart-valve reactions to dopamine-agonist treatment in Parkinson's disease. *Lancet Neurol*. England; 2007; 6: 826–9.
156. Janssen W, Schymura Y, Novoyatleva T, Kojonazarov B, Boehm M, Wietelmann A, Luitel H, Murmann K, Krompiec DR, Tretyn A, Pullamsetti SS, Weissmann N, Seeger W, et al. 5-HT_{2B} receptor antagonists inhibit fibrosis and protect from RV heart failure. *Biomed Res Int*. Hindawi Limited; 2015; 2015.
157. Andrejak M, Tribouilloy C. Drug-induced valvular heart disease: an update. *Arch Cardiovasc Dis*. Netherlands; 2013; 106: 333–9.
158. FDA (Food and Drug Administration). FDA Announces Withdrawal Fenfluramine and Dexfenfluramine (Fen-Phen) [Internet]. 1997 [cited 2017 Oct 20]. Available from: <https://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm179871.htm>
159. Fung M, Thornton A, Mybeck K, wu JHH, Hornbuckle K, Muniz E. Evaluation of the Characteristics of Safety Withdrawal of Prescription Drugs from Worldwide Pharmaceutical Markets-1960 to 1999. *Ther Innov Regul Sci*. 2001; 35: 293–317.
160. EMA (European Medicines Agency). Questions and answers on the withdrawal of medicines containing benfluorex. 2010.
161. FDA (Food and Drug Administration). FDA Drug Safety Communication: FDA Recommends Against the Continued Use of Meridia (sibutramine) [Internet]. 2010 [cited 2019 Sep 27]. Available from:

- <https://www.fda.gov/drugs/drug-safety-and-availability/fda-drug-safety-communication-fda-recommends-against-continued-use-meridia-sibutramine>
162. Yun J, Chung E, Choi KH, Cho DH, Song YJ, Han KM, Cha HJ, Shin JS, Seong WK, Kim YH, Kim HS. Cardiovascular safety pharmacology of sibutramine. *Biomol Ther. Korean Society of Applied Pharmacology*; 2015; 23: 386–9.
 163. Kim K-S, Kim E-J, Lee H-A, Park S-J. Effect of sibutramine HCl on cardiac hERG K⁺ channel. *Mol Cell Biochem* [Internet]. 2009 [cited 2019 Sep 27]; 320: 125–31. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18781376>
 164. Adler A, Viskin S, Bhuiyan ZA, Eisenberg E, Rosso R. Propoxyphene-induced torsades de pointes. *Hear Rhythm. United States*; 2011; 8: 1952–4.
 165. EMEA (European Medicines Evaluation Agency). EMEA public statement on the recommendation to suspend the marketing authorisation for ORLAAM (levacetylmethadol) in the European Union. 2001.
 166. FDA (Food and Drug Administration). FDA Drug Safety Communication: FDA recommends against the continued use of propoxyphene. [Internet]. 2010 [cited 2017 Oct 20]. Available from: <https://www.fda.gov/Drugs/DrugSafety/ucm234338.htm>
 167. FDA (Food and Drug Administration). Determination That ORLAAM (Levomethadyl Acetate Hydrochloride) Oral Solution, 10 Milligrams/Milliliter, Was Not Withdrawn From Sale for Reasons of Safety or Effectiveness. [Internet]. 2011 [cited 2017 Oct 20]. Available from: <https://www.federalregister.gov/documents/2011/06/06/2011-13884/determination-that-omlaam-levomethadyl-acetate-hydrochloride-oral-solution-10-milligramsmilliliter>
 168. FDA (Food and Drug Administration). Withdrawal of Approval of 73 New Drug Applications and 62 Abbreviated New Drug Applications. 2007.
 169. Arana A, Margulis A V., McQuay LJ, Ziemiecki R, Bartsch JL, Rothman KJ, Franks B, D'Silva M, Appenteng K, Varas-Lorenzo C, Perez-Gutthann S. Variation in Cardiovascular Risk Related to Individual

- Antimuscarinic Drugs Used to Treat Overactive Bladder: A UK Cohort Study. *Pharmacotherapy*. Pharmacotherapy Publications Inc.; 2018; 38: 628–37.
170. Thomas SH, Higham PD, Hartigan-Go K, Kamali F, Wood P, Campbell RW, Ford GA. Concentration dependent cardiotoxicity of terodiline in patients treated for urinary incontinence. *Br Heart J* [Internet]. 1995 [cited 2019 Sep 27]; 74: 53–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/7662454>
 171. Lin J, Guo J, Gang H, Wojciechowski P, Wigle JT, Zhang S. Intracellular K⁺ is required for the inactivation-induced high-affinity binding of cisapride to HERG channels. *Mol Pharmacol* [Internet]. 2005 [cited 2019 Sep 27]; 68: 855–65. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15967876>
 172. WHO (World Health Organization). Pharmaceuticals: Restrictions in Use and Availability [Internet]. 2001 [cited 2019 Sep 27]. Available from: <https://apps.who.int/medicinedocs/en/d/Js2203e/2.12.html#Js2203e.2.12>
 173. FDA (Food and Drug Administration). Drug Safety and Availability - Zelnorm (tegaserod maleate) Information. [Internet]. 2007 [cited 2017 Aug 1]. Available from: <https://www.fda.gov/Drugs/DrugSafety/ucm103223.htm>
 174. Tack J, Camilleri M, Chang L, Chey WD, Galligan JJ, Lacy BE, Müller-Lissner S, Quigley EMM, Schuurkes J, De Maeyer JH, Stanghellini V. Systematic review: cardiovascular safety profile of 5-HT(4) agonists developed for gastrointestinal disorders. *Aliment Pharmacol Ther* [Internet]. 2012 [cited 2019 Sep 27]; 35: 745–67. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22356640>
 175. Church J, Fletcher EJ, Abdel-Hamid K, MacDonald JF. Loperamide blocks high-voltage-activated calcium channels and N-methyl-D-aspartate-evoked responses in rat and mouse cultured hippocampal pyramidal neurons. *Mol Pharmacol* [Internet]. 1994 [cited 2019 Sep 27]; 45: 747–57. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/8183255>
 176. Spinner HL, Lonardo NW, Mulamalla R, Stehlik J. Ventricular

- tachycardia associated with high-dose chronic loperamide use. *Pharmacotherapy* [Internet]. 2015 [cited 2019 Sep 27]; 35: 234–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25645123>
177. Upadhyay A, Bodar V, Malekzadegan M, Singh S, Frumkin W, Mangla A, Doshi K. Loperamide Induced Life Threatening Ventricular Arrhythmia. *Case Reports Cardiol*. Hindawi Limited; 2016; 2016: 1–3.
 178. FDA (Food and Drug Administration). FDA warns about serious heart problems with high doses of the antidiarrheal medicine loperamide (Imodium), including from abuse and misuse. 2016.
 179. Juurlink DN, Dormuth CR, Huang A, Hellings C, Paterson JM, Raymond C, Kozyrskyj A, Moride Y, Macdonald EM, Mamdani MM. Proton pump inhibitors and the risk of adverse cardiac events. *PLoS One* [Internet]. 2013 [cited 2019 Sep 27]; 8: e84890. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24386430>
 180. Ghebremariam YT, LePendou P, Lee JC, Erlanson DA, Slaviero A, Shah NH, Leiper J, Cooke JP. Unexpected effect of proton pump inhibitors: elevation of the cardiovascular risk factor asymmetric dimethylarginine. *Circulation* [Internet]. 2013 [cited 2019 Sep 27]; 128: 845–53. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23825361>
 181. Hazarika S, Annex BH. Can Endothelial Injury Provide the Passage to Explain the Vascular Effects of Proton Pump Inhibitors? *Circulation Research*. Lippincott Williams and Wilkins; 2016. p. 1858–60.
 182. Gottlieb S. Antihistamine drug withdrawn by manufacturer. *BMJ* [Internet]. 1999 [cited 2019 Sep 27]; 319: 7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/10390435>
 183. Jo S-H, Hong H-K, Chong SH, Lee HS, Choe H. H(1) antihistamine drug promethazine directly blocks hERG K(+) channel. *Pharmacol Res* [Internet]. 2009 [cited 2019 Sep 27]; 60: 429–37. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19497368>
 184. Mason RP, Walter MF, Day CA, Jacob RF. A biological rationale for the cardiotoxic effects of rofecoxib: comparative analysis with other COX-2 selective agents and NSAids. *Subcell Biochem* [Internet]. 2007 [cited 2019 Sep 27]; 42: 175–90. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17612051>

185. FDA (Food and Drug Administration). Consumer Updates - FDA Strengthens Warning of Heart Attack and Stroke Risk for Non-Steroidal Anti-Inflammatory Drugs [Internet]. 2015 [cited 2017 Jun 26]. Available from:
<https://www.fda.gov/ForConsumers/ConsumerUpdates/ucm453610.htm>
186. Sondergaard KB, Weeke P, Wissenberg M, Schjerning Olsen A-M, Fosbol EL, Lippert FK, Torp-Pedersen C, Gislason GH, Folke F. Non-steroidal anti-inflammatory drug use is associated with increased risk of out-of-hospital cardiac arrest: a nationwide case-time-control study. *Eur Hear journal Cardiovasc Pharmacother* [Internet]. 2017 [cited 2019 Sep 27]; 3: 100–7. Available from:
<http://www.ncbi.nlm.nih.gov/pubmed/28025218>
187. Chandra M, Kent PJ. Cyclo-oxygenase-2 Inhibitors and Peripheral Thrombosis - A Case Report Demonstrating a Possible Adverse Effect. *EJVES Extra*. 2009; 17: 14–6.
188. Carnieto Jr. A, Dourado PMM, Luz PL da, Chagas ACP. Selective cyclooxygenase-2 inhibition protects against myocardial damage in experimental acute ischemia. *Clinics* [Internet]. 2009 [cited 2019 Sep 27]; 64: 245–52. Available from:
http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1807-59322009000300016&lng=en&nrm=iso&tlng=en
189. Krotz F, Schiele TM, Klauss V, Sohn H-Y. Selective COX-2 inhibitors and risk of myocardial infarction. *J Vasc Res* [Internet]. [cited 2019 Sep 27]; 42: 312–24. Available from:
<http://www.ncbi.nlm.nih.gov/pubmed/15976506>
190. Keserü GM. Prediction of hERG potassium channel affinity by traditional and hologram qSAR methods. *Bioorg Med Chem Lett* [Internet]. 2003 [cited 2019 Sep 27]; 13: 2773–5. Available from:
<http://www.ncbi.nlm.nih.gov/pubmed/12873512>
191. Albert RK, Schuller JL. Macrolide antibiotics and the risk of cardiac arrhythmias. *Am J Respir Crit Care Med*. United States; 2014; 189: 1173–80.
192. Wang X, Ryu D, Houtkooper RH, Auwerx J. Antibiotic use and abuse: A threat to mitochondria and chloroplasts with impact on research, health,

- and environment. *BioEssays*. John Wiley and Sons Inc.; 2015; 37: 1045–53.
193. Anderson ME, Mazur A, Yang T, Roden DM. Potassium current antagonist properties and proarrhythmic consequences of quinolone antibiotics. *J Pharmacol Exp Ther*. United States; 2001; 296: 806–10.
 194. Durante-Mangoni E, Parrella A, Vitrone M, Rago A, Pafundi PC, Nigro G, Utili R, Russo V. Electrophysiological Adverse Effects of Direct Acting Antivirals in Patients With Chronic Hepatitis C. *J Clin Pharmacol* [Internet]. 2017 [cited 2019 Sep 27]; 57: 924–30. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28117887>
 195. Haverkamp W, Breithardt G, Camm AJ, Janse MJ, Rosen MR, Antzelevitch C, Escande D, Franz M, Malik M, Moss A, Shah R. The potential for QT prolongation and pro-arrhythmia by non-anti- arrhythmic drugs: Clinical and regulatory implications: Report on a Policy Conference of the European Society of Cardiology. *Cardiovasc Res*. 2000; 47: 219–33.
 196. Katchman AN, Koerner J, Tosaka T, Woosley RL, Ebert SN. Comparative evaluation of HERG currents and QT intervals following challenge with suspected torsadogenic and nontorsadogenic drugs. *J Pharmacol Exp Ther*. United States; 2006; 316: 1098–106.
 197. Gibson JK, Yue Y, Bronson J, Palmer C, Numann R. Human stem cell-derived cardiomyocytes detect drug-mediated changes in action potentials and ion currents. *J Pharmacol Toxicol Methods*. United States; 2014; 70: 255–67.
 198. Johannesen L, Vicente J, Mason JW, Sanabria C, Waite-Labott K, Hong M, Guo P, Lin J, Sørensen JS, Galeotti L, Florian J, Ugander M, Stockbridge N, et al. Differentiating drug-induced multichannel block on the electrocardiogram: Randomized study of dofetilide, quinidine, ranolazine, and verapamil. *Clin Pharmacol Ther*. Nature Publishing Group; 2014; 96: 549–58.
 199. Ridley JM, Dooley PC, Milnes JT, Witchel HJ, Hancox JC. Lidoflazine is a high affinity blocker of the HERG K(+)channel. *J Mol Cell Cardiol* [Internet]. 2004 [cited 2019 Sep 27]; 36: 701–5. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15135665>

200. EMA (European Medicines Agency). Questions and answers on the suspension of buflomedil-containing medicines [Internet]. 2012. Available from: https://www.ema.europa.eu/en/documents/referral/questions-answers-suspension-buflomedil-containing-medicines_en.pdf
201. MHRA (Medicines and Healthcare products Regulatory Agency). Orciprenaline sulphate (Alupent): withdrawal due to unfavourable benefit-risk profile [Internet]. 2009 [cited 2017 Oct 20]. Available from: <https://www.gov.uk/drug-safety-update/orciprenaline-sulphate-alupent-withdrawal-due-to-unfavourable-benefit-risk-profile>
202. Arensberg D, Richardson DW, Obias-Manno D, Echt DS, Barker AH, Baker A, Friedman L, Peters RW, Mitchell LB, Huther ML, Liebson PR, Greene HL. Mortality and Morbidity in Patients Receiving Encainide, Flecainide, or Placebo. *N Engl J Med* [Internet]. 2010 [cited 2019 Sep 27]; 324: 781–8. Available from: <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med3&NEWS=N&AN=1713648%0Ahttp://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed5&NEWS=N&AN=21279734>
203. Kober L, Torp-Pedersen C, McMurray JJ V, Gotzsche O, Levy S, Crijns H, Amlie J, Carlsen J. Increased mortality after dronedarone therapy for severe heart failure. *N Engl J Med*. United States; 2008; 358: 2678–87.
204. Connolly SJ, Camm AJ, Halperin JL, Joyner C, Alings M, Amerena J, Atar D, Avezum Á, Blomström P, Borggrefe M, Budaj A, Chen SA, Ching CK, et al. Dronedarone in high-risk permanent atrial fibrillation. *N Engl J Med*. Massachusetts Medical Society; 2011; 365: 2268–76.
205. FDA (Food and Drug Administration). Drugs@FDA: FDA Approved Drug Products [Internet]. 2019 [cited 2019 Sep 29]. Available from: <https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm?event=overview.process&ApplNo=017535>
206. Maru S, Koch GG, Stender M, Clark D, Gibowski L, Petri H, White AD, Simpson RJ. Antidiabetic drugs and heart failure risk in patients with type 2 diabetes in the U.K. primary care setting. *Diabetes Care* [Internet]. 2005 [cited 2019 Sep 27]; 28: 20–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15616228>

207. Kannel WB, Hjortland M, Castelli WP. Role of diabetes in congestive heart failure: the Framingham study. *Am J Cardiol* [Internet]. 1974 [cited 2019 Sep 27]; 34: 29–34. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/4835750>
208. NHS. Avandia diabetes drug suspended. 2010.
209. FDA (Food and Drug Administration). FDA Drug Safety Communication: FDA eliminates the Risk Evaluation and Mitigation Strategy (REMS) for rosiglitazone-containing diabetes medicines. [Internet]. 2015 [cited 2017 Oct 20]. Available from: <https://www.fda.gov/Drugs/DrugSafety/ucm476466.htm>
210. FDA (Food and Drug Administration). Diabetes Medications Containing Saxagliptin and Alogliptin: Drug Safety Communication - Risk of Heart Failure. [Internet]. 2016 [cited 2017 Oct 20]. Available from: https://www.fda.gov/Safety/MedWatch/SafetyInformation/SafetyAlertsforHumanMedicalProducts/ucm494252.htm?source=govdelivery&utm_medium=email&utm_source=govdelivery
211. Boehringer Ingelheim GmbH. Expert statement on the non-clinical and clinical evaluation of the potential for delayed ventricular repolarization (QT Interval Prolongation) by Clobutinol (No. U07-0134). 2007.
212. Takahara A, Sasaki R, Nakamura M, Sendo A, Sakurai Y, Namekata I, Tanaka H. Clobutinol delays ventricular repolarization in the guinea pig heart: comparison with cardiac effects of HERG K⁺ channel inhibitor E-4031. *J Cardiovasc Pharmacol* [Internet]. 2009 [cited 2019 Sep 27]; 54: 552–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19770670>
213. Henninger C, Fritz G. Statins in anthracycline-induced cardiotoxicity: Rac and Rho, and the heartbreakers. *Cell Death Dis* [Internet]. 2017 [cited 2019 Sep 27]; 8: e2564. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28102848>
214. Bromage DI, Yellon DM. The pleiotropic effects of metformin: Time for prospective studies. *Cardiovasc Diabetol*. BioMed Central Ltd.; 2015; 14.
215. Lee JW, Morris JK, Wald NJ. Grapefruit Juice and Statins. *Am J Med* [Internet]. 2016 [cited 2019 Sep 27]; 129: 26–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26299317>

216. Wang Y-C, Hsieh T-C, Chou C-L, Wu J-L, Fang T-C. Risks of Adverse Events Following Coprescription of Statins and Calcium Channel Blockers: A Nationwide Population-Based Study. *Medicine (Baltimore)* [Internet]. 2016 [cited 2019 Sep 27]; 95: e2487. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26765458>
217. Arunachalam G, Upadhyay R, Ding H, Triggle CR. MicroRNA Signature and Cardiovascular Dysfunction. *J Cardiovasc Pharmacol*. United States; 2015; 65: 419–29.
218. Chistiakov DA, Orekhov AN, Bobryshev Y V. Cardiac-specific miRNA in cardiogenesis, heart function, and cardiac pathology (with focus on myocardial infarction). *J Mol Cell Cardiol* [Internet]. 2016 [cited 2019 Sep 27]; 94: 107–21. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27056419>
219. Su X, Liang H, Wang H, Chen G, Jiang H, Wu Q, Liu T, Liu Q, Yu T, Gu Y, Yang B, Shan H. Over-expression of microRNA-1 causes arrhythmia by disturbing intracellular trafficking system. *Sci Rep* [Internet]. 2017 [cited 2019 Sep 27]; 7: 46259. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28397788>
220. Xing Z, Li D, Yang L, Xi Y, Su X. MicroRNAs and anticancer drugs. *Acta Biochimica et Biophysica Sinica*. 2014. p. 233–9.
221. Rodriguez B, Carusi A, Abi-Gerges N, Ariga R, Britton O, Bub G, Bueno-Orovio A, Burton RAB, Carapella V, Cardone-Noot L, Daniels MJ, Davies MR, Dutta S, et al. Human-based approaches to pharmacology and cardiology: An interdisciplinary and intersectorial workshop. *Europace*. Oxford University Press; 2016. p. 1287–98.
222. Bloom MW, Hamo CE, Cardinale D, Ky B, Nohria A, Baer L, Skopicki H, Lenihan DJ, Gheorghiade M, Lyon AR, Butler J. Cancer Therapy-Related Cardiac Dysfunction and Heart Failure: Part 1: Definitions, Pathophysiology, Risk Factors, and Imaging. *Circ Heart Fail* [Internet]. 2016; 9: e002661–e002661. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/26747861>
223. Omotayo Fasan LD. Acute Heart Failure in a Patient with Acute Myeloid Leukemia following Daunorubicin Treatment: a Case Report. *J Leuk. OMICS Publishing Group*; 2015; 03.

224. Rateesh S, Luis SA, Luis CR, Hughes B, Nicolae M. Myocardial infarction secondary to 5-fluorouracil: not an absolute contraindication to rechallenge? *Int J Cardiol* [Internet]. 2014 [cited 2019 Sep 27]; 172: e331-3. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24447743>
225. Fukunaga T, Soejima H, Sugamura K, Kojima S, Sugiyama S, Sakamoto T, Yoshimura M, Tanoue T, Kageshita T, Ono T, Ogawa H. Acute myocardial infarction induced by cisplatin-based combination chemotherapy for malignant melanoma: a case report. *J Cardiol* [Internet]. 2006 [cited 2019 Sep 27]; 47: 191–5. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16637253>
226. Dermitzakis E V., Kimiskidis VK, Lazaridis G, Alexopoulou Z, Timotheadou E, Papanikolaou A, Romanidou O, Georgiadis G, Kalogeras KT, Tsiptsios I, Tarlatzis B, Fountzilas G. The impact of paclitaxel and carboplatin chemotherapy on the autonomous nervous system of patients with ovarian cancer. *BMC Neurol*. BioMed Central Ltd.; 2016; 16.
227. Esber C, Breathett K, Sachak T, Moore S, Lilly SM. Acute Myocardial Infarction in Patient With Triple Negative Breast Cancer After Paclitaxel Infusion: A Case Report. *Cardiol Res* [Internet]. 2014 [cited 2019 Sep 27]; 5: 108–11. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28348706>
228. Shimoyama M, Murata Y, Sumi KI, Hamazoe R, Komuro I. Docetaxel induced cardiotoxicity. *Heart* [Internet]. 2001 [cited 2019 Sep 27]; 86: 219. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/11454849>
229. Brose MS, Frenette CT, Keefe SM, Stein SM. Management of Sorafenib-related adverse events: A clinician's perspective. *Semin Oncol*. 2014; 41.
230. Moy B, Goss PE. Lapatinib-Associated Toxicity and Practical Management Recommendations. *Oncologist*. Alphamed Press; 2007; 12: 756–65.
231. Grande E, Kreissl MC, Filetti S, Newbold K, Reinisch W, Robert C, Schlumberger M, Tolstrup LK, Zamorano JL, Capdevila J. Vandetanib in advanced medullary thyroid cancer: Review of adverse event management strategies. *Advances in Therapy*. 2013. p. 945–66.

232. Rivers ZT, Oostra DR, Westholder JS, Vercellotti GM. Romidepsin-associated cardiac toxicity and ECG changes: A case report and review of the literature. *J Oncol Pharm Pract* [Internet]. 2018 [cited 2019 Sep 27]; 24: 56–62. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27708192>
233. Loewen PS. Review of the selective COX-2 inhibitors celecoxib and rofecoxib: Focus on clinical aspects. *Canadian Journal of Emergency Medicine*. Canadian Medical Association; 2002. p. 268–75.
234. Arfe A, Scotti L, Varas-Lorenzo C, Nicotra F, Zambon A, Kollhorst B, Schink T, Garbe E, Herings R, Straatman H, Schade R, Villa M, Lucchi S, et al. Non-steroidal anti-inflammatory drugs and risk of heart failure in four European countries: nested case-control study. *BMJ*. England; 2016; 354: i4857.
235. Elsaie ML. Cardiovascular collapse developing after topical anesthesia. *Dermatology* [Internet]. 2007 [cited 2019 Sep 27]; 214: 194. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17341877>
236. Kurdyak PA, Manno M, Gomes T, Mamdani MM, Juurlink DN. Antidepressants, metoprolol and the risk of bradycardia. *Ther Adv Psychopharmacol*. 2012; 2: 43–9.
237. Howell C, Wilson AD, Waring WS. Cardiovascular toxicity due to venlafaxine poisoning in adults: A review of 235 consecutive cases. *British Journal of Clinical Pharmacology*. 2007. p. 192–7.
238. Johnson EM, Whyte E, Mulsant BH, Pollock BG, Weber E, Begley AE, Reynolds CF. Cardiovascular changes associated with venlafaxine in the treatment of late-life depression. *Am J Geriatr Psychiatry* [Internet]. 2006 [cited 2019 Sep 27]; 14: 796–802. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16943176>
239. Cannistra LB, Davis SM, Bauman AG. Valvular heart disease associated with dexfenfluramine. *N Engl J Med* [Internet]. 1997 [cited 2019 Sep 27]; 337: 636. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9280831>
240. Malak J, Deprez P, Adam JF. Valve disease associated with chronic intake of fenfluramines. *Journal of Laboratory and Clinical Medicine*. Mosby Inc.; 1998. p. 475.

241. Connolly HM, Crary JL, McGoon MD, Hensrud DD, Edwards BS, Edwards WD, Schaff H V. Valvular heart disease associated with fenfluramine phentermine. *N Engl J Med*. 1997; 337: 581–8.
242. Rothman RB, Ayestas MA, Dersch CM, Baumann MH. Aminorex, fenfluramine, and chlorphentermine are serotonin transporter substrates. Implications for primary pulmonary hypertension. *Circulation* [Internet]. 1999 [cited 2019 Sep 27]; 100: 869–75. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/10458725>
243. Weill A, Païta M, Tuppin P, Fagot J-P, Neumann A, Simon D, Ricordeau P, Montastruc J-L, Allemand H. Benfluorex and valvular heart disease: a cohort study of a million people with diabetes mellitus. *Pharmacoepidemiol Drug Saf* [Internet]. 2010 [cited 2019 Sep 27]; 19: 1256–62. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20945504>
244. Eroglu E, Gemici G, Bayrak F, Kalkan AK, Degertekin M. Acute myocardial infarction in a 24 year-old man possibly associated with sibutramine use. *Int J Cardiol*. 2009; 137.
245. James WPT, Caterson ID, Coutinho W, Finer N, Van Gaal LF, Maggioni AP, Torp-Pedersen C, Sharma AM, Shepherd GM, Rode RA, Renz CL, SCOUT Investigators. Effect of sibutramine on cardiovascular outcomes in overweight and obese subjects. *N Engl J Med* [Internet]. 2010 [cited 2019 Sep 27]; 363: 905–17. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20818901>
246. Yim K-MA, Ng HW, Chan C-K, Yip G, Lau FL. Sibutramine-induced acute myocardial infarction in a young lady. *Clin Toxicol (Phila)* [Internet]. 2008 [cited 2019 Sep 27]; 46: 877–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18788006>
247. Wieneke H, Conrads H, Wolstein J, Breuckmann F, Gastpar M, Erbel R, Scherbaum N. Levo-alpha-acetylmethadol (LAAM) induced QTc-prolongation - results from a controlled clinical trial. *Eur J Med Res* [Internet]. 2009 [cited 2019 Sep 27]; 14: 7–12. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19258204>
248. Swank KA, Wu E, Kortepeter C, McAninch J, Levin RL. Adverse event detection using the FDA post-marketing drug safety surveillance

- system: Cardiotoxicity associated with loperamide abuse and misuse. *J Am Pharm Assoc* (2003) [Internet]. [cited 2019 Sep 27]; 57: S63–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28073687>
249. Upadhyay A, Bodar V, Malekzadegan M, Singh S, Frumkin W, Mangla A, Doshi K. Loperamide Induced Life Threatening Ventricular Arrhythmia. *Case Reports Cardiol* [Internet]. 2016; 2016: 1–3. Available from: <http://www.hindawi.com/journals/cric/2016/5040176/>
 250. Quigley EMM. Prokinetics in the Management of Functional Gastrointestinal Disorders. *Curr Gastroenterol Rep* [Internet]. 2017 [cited 2019 Sep 27]; 19: 53. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28887755>
 251. Stewart DA, Taylor J, Ghosh S, Macphee GJ, Abdullah I, McLenachan JM, Stott DJ. Terodiline causes polymorphic ventricular tachycardia due to reduced heart rate and prolongation of QT interval. *Eur J Clin Pharmacol* [Internet]. 1992 [cited 2019 Sep 27]; 42: 577–80. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/1623896>
 252. Bishop RO, Gaudry PL. Prolonged Q-T interval following astemizole overdose. *Arch Emerg Med. England*; 1989; 6: 63–5.
 253. Shah A, Yousuf T, Ziffra J, Zaidi A, Raghuvir R. Diphenhydramine and QT prolongation - A rare cardiac side effect of a drug used in common practice. *J Cardiol Cases. Elsevier Ltd*; 2015; 12: 126–9.
 254. Monahan BP, Ferguson CL, Killeavy ES, Lloyd BK, Troy J, Cantilena LR. Torsades de pointes occurring in association with terfenadine use. *JAMA* [Internet]. 1990 [cited 2019 Sep 27]; 264: 2788–90. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/1977935>
 255. Berni E, de Voogd H, Halcox JP, Butler CC, Bannister CA, Jenkins-Jones S, Jones B, Ouwers M, Currie CJ. Risk of cardiovascular events, arrhythmia and all-cause mortality associated with clarithromycin versus alternative antibiotics prescribed for respiratory tract infections: a retrospective cohort study. *BMJ Open. England*; 2017; 7: e013398.
 256. Wong AYS, Root A, Douglas IJ, Chui CSL, Chan EW, Ghebremichael-Weldeselassie Y, Siu CW, Smeeth L, Wong ICK. Cardiovascular outcomes associated with use of clarithromycin: Population based study. *BMJ. BMJ Publishing Group*; 2016; 352.

257. Ray WA, Murray KT, Hall K, Arbogast PG, Stein CM. Azithromycin and the risk of cardiovascular death. *N Engl J Med*. Massachusetts Medical Society; 2012; 366: 1881–90.
258. Carr A, Cooper DA. Adverse effects of antiretroviral therapy. *Lancet* (London, England). England; 2000; 356: 1423–30.
259. Shantsila E, Watson T, Lip GYH. Drug-induced QT-interval prolongation and proarrhythmic risk in the treatment of atrial arrhythmias. *Europace* [Internet]. 2007 [cited 2019 Sep 27]; 9 Suppl 4: iv37-44. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17766322>
260. Grenadier E, Keidar S, Alpan G, Marmor A, Palant A. Prenylamine-induced ventricular tachycardia and syncope controlled by ventricular pacing. *Br Heart J* [Internet]. 1980 [cited 2019 Sep 27]; 44: 330–4. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/7426190>
261. Hanley SP, Hampton JR. Ventricular arrhythmias associated with lidoflazine: side-effects observed in a randomized trial. *Eur Heart J* [Internet]. 1983 [cited 2019 Sep 27]; 4: 889–93. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/6363079>
262. Legras A, Piquemal R, Furet Y, Dequin PF, Perrotin D. Buflomedil poisoning: five cases with cardiotoxicity. *Intensive Care Med* [Internet]. 1996 [cited 2019 Sep 27]; 22: 57–61. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/8857439>
263. Nissen SE, Wolski K. Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes. *N Engl J Med*. Massachusetts Medical Society; 2007; 356: 2457–71.
264. Scirica BM, Bhatt DL, Braunwald E, Steg PG, Davidson J, Hirshberg B, Ohman P, Frederich R, Wiviott SD, Hoffman EB, Cavender MA, Udell JA, Desai NR, et al. Saxagliptin and cardiovascular outcomes in patients with type 2 diabetes mellitus. *N Engl J Med*. United States; 2013; 369: 1317–26.
265. White WB, Cannon CP, Heller SR, Nissen SE, Bergenstal RM, Bakris GL, Perez AT, Fleck PR, Mehta CR, Kupfer S, Wilson C, Cushman WC, Zannad F, et al. Alogliptin after acute coronary syndrome in patients with type 2 diabetes. *N Engl J Med* [Internet]. 2013 [cited 2019 Sep 27]; 369: 1327–35. Available from:

- <http://www.ncbi.nlm.nih.gov/pubmed/23992602>
266. Lee BKB, Tiong KH, Chang JK, Liew CS, Abdul Rahman ZA, Tan AC, Khang TF, Cheong SC. DeSigN: connecting gene expression with therapeutics for drug repurposing and development. *BMC Genomics* [Internet]. 2017 [cited 2019 Sep 29]; 18: 934. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28198666>
 267. Iorio F, Bosotti R, Scacheri E, Belcastro V, Mithbaokar P, Ferriero R, Murino L, Tagliaferri R, Brunetti-Pierri N, Isacchi A, di Bernardo D. Discovery of drug mode of action and drug repositioning from transclorio, F., Bosotti, R., Scacheri, E., Belcastro, V., Mithbaokar, P., Ferriero, R., ... di Bernardo, D. (2010). Discovery of drug mode of action and drug repositioning from transcriptional r. *Proc Natl Acad Sci* [Internet]. 2010 [cited 2019 Sep 29]; 107: 14621–6. Available from: <http://www.pnas.org/cgi/doi/10.1073/pnas.1000138107>
 268. Brown AS, Patel CJ. A standard database for drug repositioning. *Sci Data*. Nature Publishing Groups; 2017; 4.
 269. Chaudhari U, Nemade H, Wagh V, Gaspar JA, Ellis JK, Srinivasan SP, Spitkovski D, Nguemo F, Louisse J, Bremer S, Hescheler J, Keun HC, Hengstler JG, et al. Identification of genomic biomarkers for anthracycline-induced cardiotoxicity in human iPSC-derived cardiomyocytes: an in vitro repeated exposure toxicity approach for safety assessment. *Arch Toxicol* [Internet]. 2016 [cited 2019 Sep 29]; 90: 2763–77. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26537877>
 270. McSweeney KM, Bozza WP, Alterovitz WL, Zhang B. Transcriptomic profiling reveals p53 as a key regulator of doxorubicin-induced cardiotoxicity. *Cell Death Discov*. Springer Nature; 2019; 5.
 271. Reyes ME, Ma J, Grove ML, Ater JL, Morrison AC, Hildebrandt MAT. RNA sequence analysis of inducible pluripotent stem cell-derived cardiomyocytes reveals altered expression of DNA damage and cell cycle genes in response to doxorubicin. *Toxicol Appl Pharmacol*. Academic Press Inc.; 2018; 356: 44–53.
 272. Holmgren G, Synnergren J, Bogestål Y, Améen C, Åkesson K, Holmgren S, Lindahl A, Sartipy P. Identification of novel biomarkers for

- doxorubicin-induced toxicity in human cardiomyocytes derived from pluripotent stem cells. *Toxicology*. Elsevier Ireland Ltd; 2015; 328: 102–11.
273. Kopljär I, Gallacher DJ, De Bondt A, Cougnaud L, Vlamincx E, Van den Wyngaert I, Lu HR. Functional and Transcriptional Characterization of Histone Deacetylase Inhibitor-Mediated Cardiac Adverse Effects in Human Induced Pluripotent Stem Cell-Derived Cardiomyocytes. *Stem Cells Transl Med*. Wiley; 2016; 5: 602–12.
 274. Wang H, Sheehan RP, Palmer AC, Everley RA, Boswell SA, Ron-Harel N, Ringel AE, Holton KM, Jacobson CA, Erickson AR, Maliszewski L, Haigis MC, Sorger PK. Adaptation of Human iPSC-Derived Cardiomyocytes to Tyrosine Kinase Inhibitors Reduces Acute Cardiotoxicity via Metabolic Reprogramming. *Cell Syst*. Cell Press; 2019; 8: 412-426.e7.
 275. Matsa E, Burrige PW, Yu KH, Ahrens JH, Termglinchan V, Wu H, Liu C, Shukla P, Sayed N, Churko JM, Shao N, Woo NA, Chao AS, et al. Transcriptome Profiling of Patient-Specific Human iPSC-Cardiomyocytes Predicts Individual Drug Safety and Efficacy Responses In Vitro. *Cell Stem Cell*. Cell Press; 2016; 19: 311–25.
 276. Claverie JM. Computational methods for the identification of differential and coordinated gene expression. *Human Molecular Genetics*. 1999. p. 1821–32.
 277. Ben-Dor A, Friedman N, Yakhini Z. Scoring Genes for Relevance.
 278. Jeanmougin M, de Reynies A, Marisa L, Paccard C, Nuel G, Guedj M. Should we abandon the t-Test in the analysis of gene expression microarray data: A comparison of variance modeling strategies. *PLoS One*. 2010; 5: 1–9.
 279. Rapaport F, Khanin R, Liang Y, Pirun M, Krek A, Zumbo P, Mason CE, Socci ND, Betel D. Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. *Genome Biol* [Internet]. 2013 [cited 2020 Jan 2]; 14: R95. Available from: <http://genomebiology.biomedcentral.com/articles/10.1186/gb-2013-14-9-r95>
 280. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A,

- McPherson A, Szczesniak MW, Gaffney DJ, Elo LL, Zhang X, Mortazavi A. A survey of best practices for RNA-seq data analysis. *Genome Biology*. BioMed Central Ltd.; 2016.
281. Tarca AL, Romero R, Draghici S. Analysis of microarray experiments of gene expression profiling. *Am J Obstet Gynecol*. 2006; 195: 373–88.
 282. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, et al. Gene Ontology: tool for the unification of biology. *Nat Genet* [Internet]. 2000; 25: 25–9. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/pmc3037419/>
 283. Pavel AB, Sonkin D, Reddy A. Integrative modeling of multi-omics data to identify cancer drivers and infer patient-specific gene activity. *BMC Syst Biol* [Internet]. 2016 [cited 2020 Jan 2]; 10: 16. Available from: <http://www.biomedcentral.com/1752-0509/10/16>
 284. Ozerov I V, Lezhnina K V, Izumchenko E, Artemov A V, Medintsev S, Vanhaelen Q, Aliper A, Vijg J, Osipov AN, Labat I, West MD, Buzdin A, Cantor CR, et al. In silico Pathway Activation Network Decomposition Analysis (iPANDA) as a method for biomarker development. *Nat Commun* [Internet]. The Author(s); 2016; 7: 13427. Available from: <http://dx.doi.org/10.1038/ncomms13427>
 285. Mamoshina P, Vieira A, Putin E, Zhavoronkov A. Applications of deep learning in biomedicine. *Mol Pharm*. 2016; .
 286. Mincholé A, Rodriguez B. Artificial intelligence for the electrocardiogram. *Nat Med* [Internet]. 2019; 25: 22–3. Available from: <https://doi.org/10.1038/s41591-018-0306-1>
 287. Lin E, Lane H-Y. Machine learning and systems genomics approaches for multi-omics data. *Biomark Res* [Internet]. BioMed Central; 2017; 5: 2. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/28127429>
 288. Lo Y-C, Rensi SE, Torng W, Altman RB. Machine learning in chemoinformatics and drug discovery. *Drug Discov Today* [Internet]. 2018/05/08. 2018; 23: 1538–46. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/29750902>
 289. Huang LC, Wu X, Chen JY. Predicting adverse side effects of drugs. *BMC Genomics*. 2011; 12.

290. Lancaster MC, Sobie EA. Improved Prediction of Drug-Induced Torsades de Pointes Through Simulations of Dynamics and Machine Learning Algorithms. *Clin Pharmacol Ther.* United States; 2016; 100: 371–9.
291. Wang Z, Clark NR, Ma'ayan A. Drug Induced Adverse Events Prediction with the LINCS L1000 Data. *Bioinformatics* [Internet]. 2016; : btw168. Available from: <http://bioinformatics.oxfordjournals.org/content/early/2016/04/01/bioinformatics.btw168.abstract>
292. D.E. M, I.N. M, J. H, N. V, L.G. A, J.P.F. B, Messinis DE, Melas IN, Hur J, Varshney N, Alexopoulos LG, Bai JPF. Translational systems pharmacology-based predictive assessment of drug-induced cardiomyopathy. *CPT Pharmacometrics Syst Pharmacol* [Internet]. 2018 [cited 2019 Oct 11]; 7: 166–74. Available from: <http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L621433418>
293. Banerjee P, Dehnbostel FO, Preissner R. Prediction Is a Balancing Act: Importance of Sampling Methods to Balance Sensitivity and Specificity of Predictive Models Based on Imbalanced Chemical Data Sets. *Front Chem* [Internet]. 2018; 6: 362. Available from: <https://www.frontiersin.org/article/10.3389/fchem.2018.00362>
294. Wang Z, Clark NR, Ma'ayan A. Drug-induced adverse events prediction with the LINCS L1000 data. *Bioinformatics* [Internet]. 2016 [cited 2019 Oct 11]; 32: 2338–45. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27153606>
295. Saeb S, Lonini L, Jayaraman A, Mohr DC, Kording KP. The need to approximate the use-case in clinical machine learning. *Gigascience.* 2017; 6: 1–9.
296. Noble D. A modification of the Hodgkin—Huxley equations applicable to Purkinje fibre action and pacemaker potentials. *J Physiol.* 1962; 160: 317–52.
297. Niederer SA, Lumens J, Trayanova NA. Computational models in cardiology. *Nature Reviews Cardiology.* Nature Publishing Group; 2019. p. 100–11.

298. O'Hara T, Virág L, Varró A, Rudy Y. Simulation of the Undiseased Human Cardiac Ventricular Action Potential: Model Formulation and Experimental Validation. McCulloch AD, editor. PLoS Comput Biol [Internet]. 2011 [cited 2020 Jan 18]; 7: e1002061. Available from: <https://dx.plos.org/10.1371/journal.pcbi.1002061>
299. Tomek J, Bueno-Orovio A, Passini E, Zhou X, Mincholé A, Britton O, Bartolucci C, Severi S, Shrier A, Virag L, Varro A, Rodriguez B. Development, calibration, and validation of a novel human ventricular myocyte model in health, disease, and drug block. *Elife*. eLife Sciences Publications Ltd; 2019; 8.
300. Liu C, Oikonomopoulos A, Sayed N, Wu JC. Modeling human diseases with induced pluripotent stem cells: From 2D to 3D and beyond. *Dev*. Company of Biologists Ltd; 2018; 145.
301. Shi Y, Inoue H, Wu JC, Yamanaka S. Induced pluripotent stem cell technology: A decade of progress. *Nature Reviews Drug Discovery*. Nature Publishing Group; 2017. p. 115–30.
302. Paci M, Hyttinen J, Aalto-Setälä K, Severi S. Computational models of ventricular-and atrial-like human induced pluripotent stem cell derived cardiomyocytes. *Ann Biomed Eng*. Kluwer Academic Publishers; 2013; 41: 2334–48.
303. Lei CL, Wang K, Clerx M, Johnstone RH, Hortigon-Vinagre MP, Zamora V, Allan A, Smith GL, Gavaghan DJ, Mirams GR, Polonchuk L. Tailoring Mathematical Models to Stem-Cell Derived Cardiomyocyte Lines Can Improve Predictions of Drug-Induced Changes to Their Electrophysiology. *Front Physiol* [Internet]. 2017 [cited 2020 Jan 18]; 8. Available from: <http://journal.frontiersin.org/article/10.3389/fphys.2017.00986/full>
304. Paci M, Hyttinen J, Rodriguez B, Severi S. Human induced pluripotent stem cell-derived versus adult cardiomyocytes: An in silico electrophysiological study on effects of ionic current block. *Br J Pharmacol*. John Wiley and Sons Inc.; 2015; 172: 5147–60.
305. Paci M, Passini E, Severi S, Hyttinen J, Rodriguez B. Phenotypic variability in LQT3 human induced pluripotent stem cell-derived cardiomyocytes and their response to antiarrhythmic pharmacologic

- therapy: An in silico approach. *Hear Rhythm*. Elsevier B.V.; 2017; 14: 1704–12.
306. Paci M, Pölönen R-P, Cori D, Penttinen K, Aalto-Setälä K, Severi S, Hyttinen J. Automatic Optimization of an in Silico Model of Human iPSC Derived Cardiomyocytes Recapitulating Calcium Handling Abnormalities. *Front Physiol* [Internet]. Frontiers Media S.A.; 2018; 9: 709. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/29997516>
 307. Ten Tusscher KHWJ, Noble D, Noble PJ, Panfilov A V. A model for human ventricular tissue. *Am J Physiol - Hear Circ Physiol*. 2004; 286.
 308. Koivumäki JT, Korhonen T, Tavi P. Impact of Sarcoplasmic Reticulum Calcium Release on Calcium Dynamics and Action Potential Morphology in Human Atrial Myocytes: A Computational Study. McCulloch AD, editor. *PLoS Comput Biol* [Internet]. 2011 [cited 2020 Jan 18]; 7: e1001067. Available from: <http://dx.plos.org/10.1371/journal.pcbi.1001067>
 309. Passini E, Britton OJ, Lu HR, Rohrbacher J, Hermans AN, Gallacher DJ, Greig RJH, Bueno-Orovio A, Rodriguez B. Human In Silico Drug Trials Demonstrate Higher Accuracy than Animal Models in Predicting Clinical Pro-Arrhythmic Cardiotoxicity. *Front Physiol* [Internet]. 2017; 8: 1–15. Available from: <http://journal.frontiersin.org/article/10.3389/fphys.2017.00668/full>
 310. McMillan B, Gavaghan DJ, Mirams GR. Early afterdepolarisation tendency as a simulated pro-arrhythmic risk indicator. *Toxicol Res (Camb)*. England; 2017; 6: 912–21.
 311. Dutta S, Chang KC, Beattie KA, Sheng J, Tran PN, Wu WW, Wu M, Strauss DG, Colatsky T, Li Z. Optimization of an In silico Cardiac Cell Model for Proarrhythmia Risk Assessment. *Front Physiol*. Switzerland; 2017; 8: 616.
 312. Li Z, Ridder BJ, Han X, Wu WW, Sheng J, Tran PN, Wu M, Randolph A, Johnstone RH, Mirams GR, Kuryshev Y, Kramer J, Wu C, et al. Assessment of an In Silico Mechanistic Model for Proarrhythmia Risk Prediction Under the CiPA Initiative. *Clin Pharmacol Ther*. United States; 2019; 105: 466–75.
 313. Brennan T, Fink M, Rodriguez B. Multiscale modelling of drug-induced

- effects on cardiac electrophysiological activity. *Eur J Pharm Sci.* Netherlands; 2009; 36: 62–77.
314. Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* Oxford University Press (OUP); 2000; 28: 27–30.
 315. Fabregat A, Jupe S, Matthews L, Sidiropoulos K, Gillespie M, Garapati P, Haw R, Jassal B, K€orninger F, May B, Milacic M, Roca CD, Rothfels K, et al. The Reactome Pathway Knowledgebase. *Nucleic Acids Res* [Internet]. 2018 [cited 2019 Oct 11]; 46: D649–55. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29145629>
 316. Ezkurdia I, Juan D, Rodriguez JM, Frankish A, Diekhans M, Harrow J, Vazquez J, Valencia A, Tress ML. Multiple evidence strands suggest that there may be as few as 19 000 human protein-coding genes. *Hum Mol Genet.* Oxford University Press; 2014; 23: 5866–78.
 317. Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, Gould J, Davis JF, Tubelli AA, Asiedu JK, Lahr DL, Hirschman JE, Liu Z, et al. A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell.* Cell Press; 2017; 171: 1437-1452.e17.
 318. Gillet J-P, Varma S, Gottesman MM. The Clinical Relevance of Cancer Cell Lines. *JNCI J Natl Cancer Inst* [Internet]. 2013 [cited 2020 Jan 2]; 105: 452–8. Available from: <https://academic.oup.com/jnci/article-lookup/doi/10.1093/jnci/djt007>
 319. Kuhn M, Letunic I, Jensen LJ, Bork P. The SIDER database of drugs and side effects. *Nucleic Acids Res* [Internet]. 2016 [cited 2019 Oct 11]; 44: D1075-9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26481350>
 320. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, Sajed T, Johnson D, Li C, Sayeeda Z, Assempour N, Iynkkaran I, Liu Y, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res* [Internet]. 2018 [cited 2020 Jan 12]; 46: D1074–82. Available from: <http://academic.oup.com/nar/article/46/D1/D1074/4602867>
 321. Guha R, Charlop-Powers Z, Schymanski E. Package “rcdk.” 2018.
 322. Zhang C, Zhou Y, Gu S, Wu Z, Wu W, Liu C, Wang K, Liu G, Li W, Lee PW, Tang Y. In silico prediction of hERG potassium channel blockage

- by chemical category approaches. *Toxicol Res* [Internet]. The Royal Society of Chemistry; 2016; 5: 570–82. Available from: <http://dx.doi.org/10.1039/C5TX00294J>
323. Novoselova N, Wang J, Pessler F, Maintainer FK. Feature Selection and Classification with the Embedded Validation Procedures for Biomedical Data Analysis. 2018.
 324. Lagani V, Athineou G, Farcomeni A, Tsagris M, Tsamardinos I. Feature Selection with the R Package MXM : Discovering Statistically Equivalent Feature Subsets . *J Stat Softw. Foundation for Open Access Statistic*; 2017; 80.
 325. Cohen J. A Coefficient of Agreement for Nominal Scales. *Educ Psychol Meas* [Internet]. 1960 [cited 2019 Oct 11]; 20: 37–46. Available from: <http://journals.sagepub.com/doi/10.1177/001316446002000104>
 326. Krawczyk B. Learning from imbalanced data: open challenges and future directions. *Progress in Artificial Intelligence*. Springer Verlag; 2016. p. 221–32.
 327. Kendir C, van den Akker M, Vos R, Metsemakers J. Cardiovascular disease patients have increased risk for comorbidity: A cross-sectional study in the Netherlands. *Eur J Gen Pract* [Internet]. 2018 [cited 2019 Oct 11]; 24: 45–50. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29168400>
 328. Zhang P, Wang F, Hu J, Sorrentino R. Exploring the relationship between drug side-effects and therapeutic indications. *AMIA Annu Symp Proc*. 2013; 2013: 1568–77.
 329. Lawson CA, Solis-Trapala I, Dahlstrom U, Mamas M, Jaarsma T, Kadam UT, Stromberg A. Comorbidity health pathways in heart failure patients: A sequences-of-regressions analysis using cross-sectional data from 10,575 patients in the Swedish Heart Failure Registry. *PLoS Med* [Internet]. 2018 [cited 2019 Oct 11]; 15: e1002540. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29584734>
 330. Naser N, Dilic M, Durak A, Kulic M, Pepic E, Smajic E, Kusljugic Z. The Impact of Risk Factors and Comorbidities on The Incidence of Atrial Fibrillation. *Mater Socio Medica*. *ScopeMed International Medical Journal Management and Indexing System*; 2017; 29: 231.

331. D'Amato G, Luxán G, de la Pompa JL. Notch signalling in ventricular chamber development and cardiomyopathy. *FEBS J* [Internet]. 2016 [cited 2019 Oct 11]; 283: 4223–37. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27260948>
332. Troncoso R, Ibarra C, Vicencio JM, Jaimovich E, Lavandero S. New insights into IGF-1 signaling in the heart. *Trends Endocrinol Metab* [Internet]. 2014 [cited 2019 Oct 11]; 25: 128–37. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24380833>
333. Chawla N V., Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling Technique. *J Artif Intell Res* [Internet]. 2002 [cited 2019 Oct 11]; 16: 321–57. Available from: <http://arxiv.org/abs/1106.1813>
334. Senge R, del Coz JJ, Hüllermeier E. Rectifying Classifier Chains for Multi-Label Classification. *arXiv* [Internet]. 2019 [cited 2019 Oct 11]; . Available from: <http://arxiv.org/abs/1906.02915>
335. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw* [Internet]. 2010 [cited 2019 Oct 11]; 33: 1–22. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20808728>
336. Santosa F, Symes WW. Linear Inversion of Band-Limited Reflection Seismograms. *SIAM J Sci Stat Comput. Society for Industrial & Applied Mathematics (SIAM)*; 1986; 7: 1307–30.
337. Tikhonov AN (Aleksandr N, Leonov AS (Aleksandr S, Yagola AG. Nonlinear ill-posed problems. 387 p.
338. Wright MN, Ziegler A. ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *J Stat Softw* [Internet]. American Statistical Association; 2017 [cited 2019 Oct 11]; 77. Available from: <http://www.jstatsoft.org/v77/i01/>
339. Ho TK. Random Decision Forest. *Proceedings of the 3rd International Conference on Document Analysis and Recognition. Montreal, Canada*; 1995. p. 278–82.
340. Greenwell B, Boehmke B, Cunningham J, GBM Developers. *Generalized Boosted Models: A guide to the gbm package*. 2007.
341. Breiman L. Arcing the Edge [Internet]. 1997. Available from:

- <https://statistics.berkeley.edu/tech-reports/486>
342. Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A. CatBoost: unbiased boosting with categorical features. 32nd Conference of Neural Information Processing Systems (NeurIPS 2018) [Internet]. Montreal, Canada; 2018 [cited 2019 Oct 11]. Available from: <https://github.com/catboost/catboost>
 343. Minasny B, McBratney AB. A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Comput Geosci* [Internet]. 2006 [cited 2019 Oct 11]; 32: 1378–88. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S009830040500292X>
 344. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, Müller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* [Internet]. 2011 [cited 2019 Oct 11]; 12: 77. Available from: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-12-77>
 345. Aliper A, Plis S, Artemov A, Ulloa A, Mamoshina P, Zhavoronkov A. Deep Learning Applications for Predicting Pharmacological Properties of Drugs and Drug Repurposing Using Transcriptomic Data. *Mol Pharm* [Internet]. 2016; 13: 2524–30. Available from: <http://dx.doi.org/10.1021/acs.molpharmaceut.6b00248>
 346. Xie L, He S, Yang S, Feng B, Wan K, Zhang Z, Bo X, Ding Y. Domain-Adversarial Multi-Task Framework for Novel Therapeutic Property Prediction of Compounds. *arXiv* [Internet]. 2018 [cited 2019 Oct 11]; . Available from: <http://arxiv.org/abs/1810.00867>
 347. Lamberti M, Giovane G, Garzillo EM, Avino F, Feola A, Porto S, Tombolini V, Di Domenico M. Animal models in studies of cardiotoxicity side effects from antitubercular drugs in patients and occupational exposed workers. *Biomed Res Int* [Internet]. 2014 [cited 2019 Oct 11]; 2014: 240642. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24701565>
 348. Musa A, Ghorraie LS, Zhang S-D, Glazko G, Yli-Harja O, Dehmer M, Haibe-Kains B, Emmert-Streib F. A review of connectivity map and computational approaches in pharmacogenomics. *Brief Bioinform* [Internet]. 2018 [cited 2019 Oct 11]; 19: 506–23. Available from:

- <http://www.ncbi.nlm.nih.gov/pubmed/28069634>
349. Han HW, Hahn S, Jeong HY, Jee JH, Nam MO, Kim HK, Lee DH, Lee SY, Choi DK, Yu JH, Min SH, Yoo J. LINCS L1000 dataset-based repositioning of CGP-60474 as a highly potent anti-endotoxemic agent. *Sci Rep. Nature Publishing Group*; 2018; 8.
 350. Ramsay RR, Popovic-Nikolic MR, Nikolic K, Uliassi E, Bolognesi ML. A perspective on multi-target drug discovery and design for complex diseases. *Clin Transl Med [Internet]*. 2018 [cited 2019 Oct 11]; 7: 3. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29340951>
 351. Passini E, Britton OJ, Lu HR, Rohrbacher J, Hermans AN, Gallacher DJ, Greig RJH, Bueno-Orovio A, Rodriguez B. Human In Silico Drug Trials Demonstrate Higher Accuracy than Animal Models in Predicting Clinical Pro-Arrhythmic Cardiotoxicity. *Front Physiol. Switzerland*; 2017; 8: 668.
 352. Lee K-M, Tarn W-Y. Coupling pre-mRNA processing to transcription on the RNA factory assembly line. *RNA Biol. United States*; 2013; 10: 380–90.
 353. MATLAB. version 9.2.0 (R2017a). Natick, Massachusetts: The MathWorks Inc.; 2017.
 354. The MathWorks Inc. MATLAB and Statistic Toolbox Release 2017a. Natick, Massachusetts; 2017.
 355. R Core Team. R: A language and environment for statistical computing. [Internet]. Vienna, Austria.; 2018. Available from: <https://www.r-project.org/>
 356. Soetaert K, Petzoldt T, Setzer RW. Solving Differential Equations in R: Package deSolve. *J Stat Software*; Vol 1, Issue 9 [Internet]. 2010; . Available from: <https://www.jstatsoft.org/v033/i09>
 357. Wickham H. ggplot2: Elegant Graphics for Data Analysis [Internet]. Springer-Verlag New York; 2016. Available from: <https://ggplot2.tidyverse.org>
 358. Ficker E, Dennis AT, Wang L, Brown AM. Role of the cytosolic chaperones Hsp70 and Hsp90 in maturation of the cardiac potassium channel HERG. *Circ Res. United States*; 2003; 92: e87-100.
 359. Ke Y, Ng CA, Hunter MJ, Mann SA, Heide J, Hill AP, Vandenberg JL.

- Trafficking defects in PAS domain mutant K^v11.1 channels: roles of reduced domain stability and altered domain–domain interactions. *Biochem J* [Internet]. 2013; 454: 69 LP – 77. Available from: <http://www.biochemj.org/content/454/1/69.abstract>
360. Li K, Jiang Q, Bai X, Yang Y-F, Ruan M-Y, Cai S-Q. Tetrameric Assembly of K(+) Channels Requires ER-Located Chaperone Proteins. *Mol Cell*. United States; 2017; 65: 52–65.
 361. Britton OJ, Bueno-Orovio A, Van Ammel K, Lu HR, Towart R, Gallacher DJ, Rodriguez B. Experimentally calibrated population of models predicts and explains intersubject variability in cardiac cellular electrophysiology. *Proc Natl Acad Sci U S A*. United States; 2013; 110: E2098-105.
 362. Schmidt-Glenewinkel H, Reinze E, Eils R, Brady NR. Systems biological analysis of epidermal growth factor receptor internalization dynamics for altered receptor levels. *J Biol Chem*. 2009; 284: 17243–52.
 363. Becker V, Schilling M, Bachmann J, Baumann U, Raue A, Maiwald T, Timmer J, Klingmüller U. Covering a broad dynamic range: Information processing at the erythropoietin receptor. *Science* (80-). 2010; 328: 1404–8.
 364. Wattis JAD, O'Malley B, Blackburn H, Pickersgill L, Panovska J, Byrne HM, Jackson KG. Mathematical model for low density lipoprotein (LDL) endocytosis by hepatocytes. *Bull Math Biol*. 2008; 70: 2303–33.
 365. Smith JL, Anderson CL, Burgess DE, Elayi CS, January CT, Delisle BP. Molecular pathogenesis of long QT syndrome type 2. *J Arrhythmia* [Internet]. 2016; 32: 373–80. Available from: <http://www.sciencedirect.com/science/article/pii/S1880427615001726>
 366. Hall AR, Anderson CL, Smith JL, Mirshahi T, Elayi CS, January CT, Delisle BP. Visualizing Mutation-Specific Differences in the Trafficking-Deficient Phenotype of Kv11.1 Proteins Linked to Long QT Syndrome Type 2. *Front Physiol* [Internet]. 2018; 9: 584. Available from: <https://www.frontiersin.org/article/10.3389/fphys.2018.00584>
 367. Kanner SA, Jain A, Colecraft HM. Development of a High-Throughput Flow Cytometry Assay to Monitor Defective Trafficking and Rescue of Long QT2 Mutant hERG Channels. *Front Physiol*. 2018; 9: 397.

368. Mehta A, Sequiera GL, Ramachandra CJA, Sudibyo Y, Chung Y, Sheng J, Wong KY, Tan TH, Wong P, Liew R, Shim W. Re-trafficking of hERG reverses long QT syndrome 2 phenotype in human iPS-derived cardiomyocytes. *Cardiovasc Res*. England; 2014; 102: 497–506.
369. Bertalovitz AC, Badhey MLO, McDonald T V. Synonymous nucleotide modification of the KCNH2 gene affects both mRNA characteristics and translation of the encoded hERG ion channel. *J Biol Chem*. United States; 2018; 293: 12120–36.
370. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* [Internet]. 2015 [cited 2020 Jan 13]; 43: e47–e47. Available from: <http://academic.oup.com/nar/article/43/7/e47/2414268/limma-powers-differential-expression-analyses-for>
371. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* [Internet]. 2014 [cited 2020 Jan 13]; 15: 550. Available from: <http://genomebiology.biomedcentral.com/articles/10.1186/s13059-014-0550-8>
372. Csardi G, Nepusz T. The igraph software package for complex network research. *InterJournal* [Internet]. 2006; Complex Sy: 1695. Available from: <http://igraph.org>
373. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003; 13: 2498–504.
374. Marchand A, Tomkiewicz C, Magne L, Barouki R, Garlatti M. Endoplasmic reticulum stress induction of insulin-like growth factor-binding protein-1 involves ATF4. *J Biol Chem*. 2006; 281: 19124–33.
375. Rutherford E, Mark P. What happens to the heart in chronic kidney disease. *J R Coll Physicians Edinb* [Internet]. 2017 [cited 2020 Jan 11]; 47: 76–82. Available from: http://www.rcpe.ac.uk/sites/default/files/jrcpe_47_1_mark.pdf
376. Perrone-Filardi P, Paolillo S, Costanzo P, Savarese G, Trimarco B,

- Bonow RO. The role of metabolic syndrome in heart failure. *Eur Heart J* [Internet]. 2015 [cited 2020 Jan 11]; 36: 2630–4. Available from: <https://academic.oup.com/eurheartj/article-lookup/doi/10.1093/eurheartj/ehv350>
377. Rao MS, Van Vleet TR, Ciurlionis R, Buck WR, Mittelstadt SW, Blomme EAG, Liguori MJ. Comparison of RNA-Seq and microarray gene expression platforms for the toxicogenomic evaluation of liver from short-term rat toxicity studies. *Front Genet. Frontiers Media S.A.*; 2019; 10.
 378. Stiehl DP, Tritto E, Chibout S-D, Cordier A, Moulin P. The Utility of Gene Expression Profiling from Tissue Samples to Support Drug Safety Assessments. *ILAR J* [Internet]. 2017 [cited 2020 Jan 11]; 58: 69–79. Available from: <http://academic.oup.com/ilarjournal/article/58/1/69/3859294/The-Utility-of-Gene-Expression-Profilng-from>
 379. Ivanenkov YA, Zagribelnyy BA, Aladinskiy VA. Are We Opening the Door to a New Era of Medicinal Chemistry or Being Collapsed to a Chemical Singularity? *J Med Chem* [Internet]. 2019 [cited 2020 Jan 11]; 62: 10026–43. Available from: <https://pubs.acs.org/doi/abs/10.1021/acs.jmedchem.9b00004>
 380. Roboz GJ, Ritchie EK, Carlin RF, Samuel M, Gale L, Provenzano-Gober JL, Curcio TJ, Feldman EJ, Kligfield PD. Prevalence, management, and clinical consequences of QT interval prolongation during treatment with arsenic trioxide. *J Clin Oncol. American Society of Clinical Oncology*; 2014; 32: 3723–8.
 381. Ruan Y, Liu N, Napolitano C, Priori SG. Therapeutic Strategies for Long-QT Syndrome. *Circ Arrhythmia Electrophysiol*. 2008; 1: 290–7.
 382. Anderson CL, Kuzmicki CE, Childs RR, Hintz CJ, Delisle BP, January CT. Large-scale mutational analysis of Kv11.1 reveals molecular insights into type 2 long QT syndrome. *Nat Commun* [Internet]. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.; 2014; 5: 5535. Available from: <https://doi.org/10.1038/ncomms6535>
 383. Balse E, Boycott HE. Ion Channel Trafficking: Control of Ion Channel

- Density as a Target for Arrhythmias? *Front Physiol* [Internet]. 2017 [cited 2020 Jan 11]; 8. Available from: <http://journal.frontiersin.org/article/10.3389/fphys.2017.00808/full>
384. Aromolaran AS, Boutjdir M. Cardiac ion channel regulation in obesity and the metabolic syndrome: Relevance to long QT syndrome and atrial fibrillation. *Frontiers in Physiology*. Frontiers Media S.A.; 2017.
385. Brundel BJJM, Van Gelder IC, Henning RH, Tuinenburg AE, Wietes M, Grandjean JG, Wilde AAM, Van Gilst WH, Crijns HJGM. Alterations in potassium channel gene expression in atria of patients with persistent and paroxysmal atrial fibrillation: Differential regulation of protein and mRNA levels for K⁺ channels. *J Am Coll Cardiol*. 2001; 37: 926–32.
386. Schumacher SM, Martens JR. Ion channel trafficking: A new therapeutic horizon for atrial fibrillation. *Hear Rhythm*. 2010; 7: 1309–15.
387. Balse E, Boycott HE. Ion channel trafficking: Control of ion channel density as a target for arrhythmias? *Frontiers in Physiology*. Frontiers Media S.A.; 2017.