



# Selection of safe artemisinin derivatives using a machine learning-based cardiotoxicity platform and in vitro and in vivo validation

Onat Kadioglu<sup>1</sup> · Sabine M. Klauck<sup>2</sup> · Edmond Fleischer<sup>3</sup> · Letian Shan<sup>4</sup> · Thomas Efferth<sup>1</sup>

Received: 18 March 2021 / Accepted: 21 April 2021 / Published online: 22 May 2021  
© The Author(s) 2021

## Abstract

The majority of drug candidates fails the approval phase due to unwanted toxicities and side effects. Establishment of an effective toxicity prediction platform is of utmost importance, to increase the efficiency of the drug discovery process. For this purpose, we developed a toxicity prediction platform with machine-learning strategies. Cardiotoxicity prediction was performed by establishing a model with five parameters (arrhythmia, cardiac failure, heart block, hypertension, myocardial infarction) and additional toxicity predictions such as hepatotoxicity, reproductive toxicity, mutagenicity, and tumorigenicity are performed by using Data Warrior and Pro-Tox-II software. As a case study, we selected artemisinin derivatives to evaluate the platform and to provide a list of safe artemisinin derivatives. Artemisinin from *Artemisia annua* was described first as an anti-malarial compound and later its anticancer properties were discovered. Here, random forest feature selection algorithm was used for the establishment of cardiotoxicity models. High AUC scores above 0.830 were achieved for all five cardiotoxicity indications. Using a chemical library of 374 artemisinin derivatives as a case study, 7 compounds (deoxydihydro-artemisinin, 3-hydroxy-deoxy-dihydroartemisinin, 3-desoxy-dihydroartemisinin, dihydroartemisinin-furano acetate-d3, deoxyartemisinin, artemisinin G, artemisinin B) passed the toxicity filtering process for hepatotoxicity, mutagenicity, tumorigenicity, and reproductive toxicity in addition to cardiotoxicity. Experimental validation with the cardiomyocyte cell line AC16 supported the findings from the in silico cardiotoxicity model predictions. Transcriptomic profiling of AC16 cells upon artemisinin B treatment revealed a similar gene expression profile as that of the control compound, dexrazoxane. In vivo experiments with a Zebrafish model further substantiated the in silico and in vitro data, as only slight cardiotoxicity in picomolar range was observed. In conclusion, our machine-learning approach combined with in vitro and in vivo experimentation represents a suitable method to predict cardiotoxicity of drug candidates.

**Keywords** Artificial intelligence · Cardiotoxicity · Drug discovery · Machine learning

## Abbreviations

AUC	Area under the curve
hERG	Human ether-a-go-go related gene
kNN	K-nearest neighboring
RF	Random forest
ROC	Receiver operating characteristic
SVM	Support vector machine

## Introduction

The drug discovery process comprises numerous and extensive preclinical analyses including in silico, in vitro, and in vivo experiments, before clinical trials with human subjects can be initiated. A main reason, why compounds do not reach clinical phases is the occurrence of

✉ Thomas Efferth  
efferth@uni-mainz.de

<sup>1</sup> Department of Pharmaceutical Biology, Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University, Staudinger Weg 5, 55128 Mainz, Germany

<sup>2</sup> Division of Cancer Genome Research, German Cancer Research Center (DKFZ), German Cancer Consortium (DKTK), National Center for Tumor Diseases (NCT), Heidelberg, Germany

<sup>3</sup> Fischer Organics GmbH, Weiler, Germany

<sup>4</sup> The First Affiliated Hospital, Zhejiang Chinese Medical University, Hangzhou, China

non-tolerable toxicities. Besides hepatotoxicity, hematotoxicity, tumorigenicity and cardiotoxicity may pose severe and even life-threatening side effects (Cheng et al. 2017; Lee 2003; McGowan et al. 2017; Oliveira et al. 2007). Most drug candidates drop out from the development pipeline because of these toxicities. One important criterion for the screening of novel drug candidates is the absence of cardiotoxic effects (Lee et al. 2019). The selection of potentially active compounds from large databases requires *in silico* methods as rapid and cost-effective pre-selection step for subsequent *in vitro* experiments. Coupling suitable *in silico* analyses with toxicity predictions increase the efficiency of preclinical drug development (Briggs et al. 2012; Issa et al. 2017). The selection of both active and safe compounds represents a most critical step. With emerging new concepts in artificial intelligence and machine learning, the prediction of toxicity may become not only more time- and cost-effective but also more reasonable and reliable.

Many natural compounds and derivatives thereof have proven their effectiveness in drug therapy during the past decades. About 60% of the clinically established drugs are indeed in one way or another of natural origin (Newman and Cragg 2007). Available natural compound databases serve as an invaluable source to identify novel compounds that could possess activity against certain diseases or disorders by focusing on particular target proteins.

The sesquiterpenoid artemisinin was initially described as anti-malarial compound isolated from the plant *Artemisia annua* L. Later on, it became apparent that artemisinin and its derivatives are also active against cancer, viral infections, schistosomiasis and trypanosomiasis (Efferth 2017a, b; Efferth et al. 2002, 2008; Michaelsen et al. 2015; Mu and Wang 2018; Ooko et al. 2015; Shi et al. 2018). Hence, this class of compounds seems to be suited for further drug development. This also implies that toxicity evaluation of artemisinin derivatives is a mandatory critical step for the development of second-generation artemisinin derivatives.

In the present study, we developed a cardiotoxicity prediction model focusing on five of the most important parameters (arrhythmia, cardiac failure, heart block, hypertension, myocardial infarction). In addition, hepatotoxicity, mutagenicity, tumorigenicity, and reproductive toxicity were also evaluated with corresponding *in silico* tools. With these computational tools at hand, we selected a chemical library of 374 artemisinin derivatives to evaluate the toxicity and provide a panel of presumably safe artemisinin derivatives. The *in silico* results for these compounds have been verified *in vitro* with a cytotoxicity assay using human cardiomyocytes, microarray to analyze gene expression profiles and *in vivo* cardiotoxicity assessment on a zebrafish model.

## Material and methods

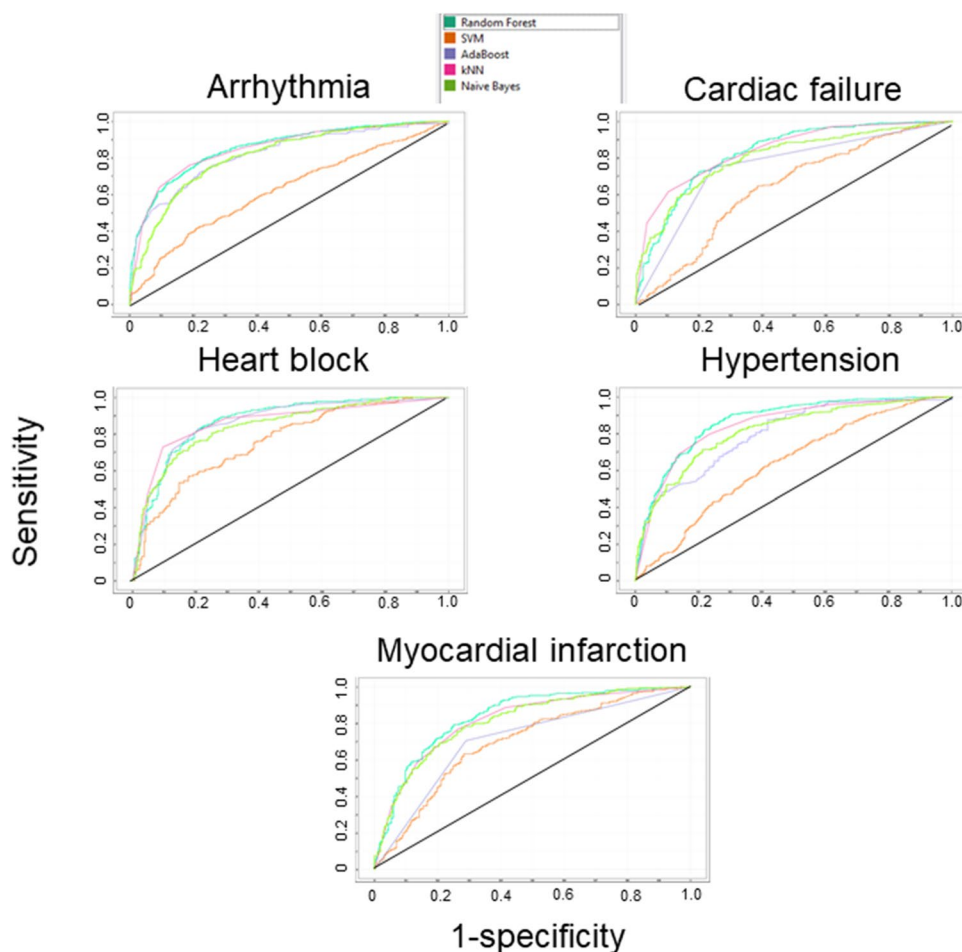
### Preparation of a library of artemisinin derivatives and calculation of chemical descriptors

The chemical structures of 374 artemisinin derivatives were retrieved from PubMed and PubChem. The Data Warrior software was used to calculate the chemical descriptors for cardiotoxicity in addition to mutagenicity, tumorigenicity, reproductive toxicity (Lopez-Lopez et al. 2019; Sander et al. 2015). After chemical descriptor calculations, correlation coefficients were calculated with SPSS software (IBM, USA). Descriptors having pairwise correlation with the cardiotoxicity parameters above 0.1 were considered to predict drug safety. For descriptors having a pairwise correlation coefficient higher than 0.9, the one with lower correlation coefficient with the cardiotoxicity parameters were excluded (Cai et al. 2018). By this strategy, relevant descriptors without the potential issue of over-fitting issue were selected to build the model.

### Cardiotoxicity prediction model establishment

The establishment of a prediction model for cardiotoxicity was performed by using the machine-learning software Orange (Ljubljana, Slovenia) (Demsar et al. 2013; Kadioglu and Efferth 2019) and the cardiotoxicity training subset of compounds deposited in the DrugBank database (<https://www.drugbank.ca/>). A total of 1451 compounds for arrhythmia, 626 compounds for cardiac failure, 545 compounds for heart block, 1163 compounds for hypertension, and 639 compounds for myocardial infarction were selected for the learning steps and model establishment. For further external validation and evaluation of the models, additional 143 compounds for arrhythmia, 538 compounds for cardiac failure, 403 compounds for heart block, 291 compounds for hypertension, and 179 compounds for myocardial infarction were selected. After applying the descriptor selection criteria by considering appropriate relevance and over-fitting issues, “logP”, “drug likeness”, “amines”, “ligand efficiency (LE)”, “alkyl-amines”, “aromatic nitrogens” and “basic nitrogens” were considered for model preparation. Various classification algorithms with the “leave-one-out” sampling method were tested, i.e. Ada Boost, k-nearest neighboring (kNN), Naive Bayes, random forest (RF), and support vector machine (SVM). Receiver operating characteristic (ROC) curves are depicted in Fig. 1. Based on this training set of known cardiotoxic compounds, the further evaluation of the established cardiotoxicity models was performed by using a list taken from the literature (Mladenka et al. 2018). The RF algorithm

**Fig. 1** Receiver operating characteristic (ROC) curves of Ada Boost, kNN, Naive Bayes, RF, SVM classification algorithms based on leave-one-out sampling for cardiotoxicity assessment models



performed better than the other classification algorithms. Overall performance for the established models is summarized in Table 1.

### Prediction of other toxicities

The prediction of hepatotoxicity and immunotoxicity was performed with the Pro-Tox-II web server (Banerjee et al. 2018). Additional toxicity parameters (mutagenicity, tumorigenicity, reproductive toxicity) were evaluated with Data Warrior software.

### Verification of cardiotoxicity predictions

At first, molecular docking on hERG (human ether-a-go-go related gene, alias *KCNH2*) protein (PDB ID: 5VA2) was performed to assess the interaction between the selected compounds and the well-known cardiotoxicity marker protein hERG. Doxorubicin is a well-known cardiotoxic drug, which was used as positive control for cardiotoxicity. Dextrazoxane is a safe cardiovascular agent, which has protective effects towards cardiotoxicity caused by anthracyclines (Liesse et al. 2018; Narayan et al. 2019; Zhang et al. 2015).

**Table 1** Performance of the in silico cardiotoxicity models based on the random forest classifier algorithm

Cardiotoxicity models	Machine learning					External validation set			
	AUC	sensitivity	specificity	overall predictive accuracy	precision	sensitivity	specificity	overall predictive accuracy	precision
Arrhythmia	0.849	0.775	0.764	0.770	0.767	0.887	0.789	0.838	0.808
Cardiac failure	0.831	0.768	0.742	0.755	0.752	0.804	0.730	0.767	0.751
Heart block	0.869	0.783	0.779	0.781	0.780	0.881	0.806	0.843	0.819
Hypertension	0.854	0.795	0.766	0.781	0.773	0.862	0.676	0.769	0.727
Myocardial infarction	0.834	0.765	0.759	0.762	0.760	0.820	0.753	0.787	0.768

This drug was utilized as negative control for cardiotoxicity. Briefly, the AutoDock 4 algorithm (Morris et al. 2009) was used for the molecular docking calculations with 2,500,000 energy evaluations and 250 runs. The AutoDock Tools 1.5.7 software (Morris et al. 2009) was applied for visualization and preparation of ligand and protein structures. The artemisinin compounds selected by the machine learning-based RF algorithm described above have been subjected to an automated and comprising molecular docking campaign by using the high-performance supercomputer MOGON (Johannes Gutenberg University, Mainz). Compound flexibilities were taken into account and a rigid receptor structure was used. The binding site of hERG was retrieved from the literature (Wang and MacKinnon 2017). Three independent defined docking runs were performed by covering the binding site on hERG. Weak binding with hERG implies safe drug profiles (Bhat and Houghton 2018; Lee et al. 2019).

### Cytotoxicity assay

To verify the cardiotoxicity of the selected compounds further, the cardiomyocyte cell line AC16 was used, which was kindly provided by Dr. James Spiers (Trinity College Dublin, Department of Pharmacology and Therapeutics). It was cultured as previously described (Elgenaidi and Spiers 2019). The cytotoxicity of the selected compounds (provided by one of the authors, E.F.) and of doxorubicin (provided by University of Mainz Clinical Pharmacy as positive control drug) towards AC16 cardiomyocytes was evaluated by the resazurin assay with three independent experiments per compounds and each 6 parallel measurements per experiment as previously described (Ooko et al. 2016).

### Microarray analyses

Microarray experiments for artemisinin B, doxorubicin and dexrazoxane were performed with 10  $\mu\text{M}$  concentration for artemisinin B and dexrazoxane on AC16 cells for 24 h, whereas the IC<sub>50</sub> concentration was used for doxorubicin (1.918  $\mu\text{M}$ ). RNA extraction and quality control were performed as previously described (Yan et al. 2020). Microarray analyses using the Affymetrix Clariom S assay (Affymetrix, Santa Clara, CA, USA) chips according to the manufacturer protocol were conducted by the Genomics and Proteomics Core Facility at the German Cancer Research Center (DKFZ, Heidelberg). Statistical analysis was performed with the Chipster software (Helsinki, Finland).

### Zebrafish handling and acute toxicity assay

Wild-type AB strain of adult zebrafish was purchased from the China Zebrafish Resource Center, Institute of Hydrobiology, China Academy of Science (Wuhan, China) and

accredited by the Association for Assessment and Accreditation of Laboratory Animal Care International (SYXX 2012–0171). Zebrafish larvae 48 h post fertilization (hpf) were obtained by natural pair-mating and housed in a light-controlled aquaculture facility with a standard 14:10 h day/night photoperiod and fed with live brine shrimp twice a day and fry flakes once a day.

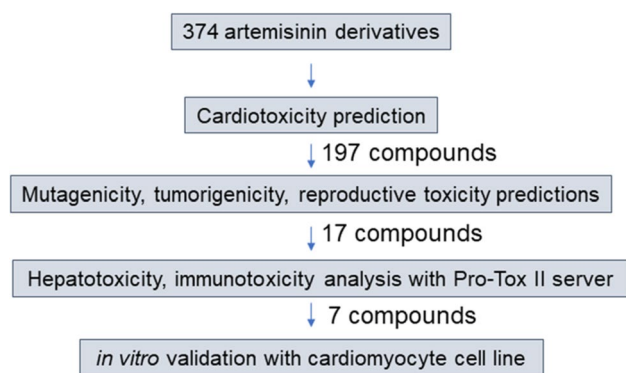
Totally, 210 zebrafish larvae 48 hpf were employed and divided into 7 groups in 6-well plates (Nest Biotech, China) with 30 fishes in each group. The basic procedure of zebrafish handling and drug treatment was recently described by us (Zheng et al. 2020). Artemisinin B was dissolved with DMSO (Sigma Aldrich, Germany) for treatment groups, and DMSO only was used for DMSO group. Artemisinin B at 0, 0.0625, 0.125, 0.25, 0.5, and 1.0 pg/ml were respectively microinjected into the yolk sac of each zebrafish larvae. All fishes were subject to visual observation and image acquisition under a dissecting stereomicroscope (Olympus Ltd, Tokyo, Japan). The death number and adverse event of zebrafishes were recorded for each group.

## Results

### Toxicity predictions

We developed a machine learning-based toxicity prediction platform focusing on cardiotoxicity, hepatotoxicity, immunotoxicity, mutagenicity, reproductive toxicity and tumorigenicity. First, we generated a cardiotoxicity prediction model with a leave-one-out sampling based random forest algorithm for five main parameters of cardiotoxicity (arrhythmia, cardiac failure, heart block, hypertension, myocardial infarction) with high-performance scores for both learning and external validation sets. The sensitivity for the learning set was in the range 0.765 to 0.795 and for the external validation set between 0.804 and 0.887 (Table 1). The specificity, overall predictive accuracy and precision for the learning and external validation set ranged between 0.676 and 0.843 (Table 1).

Secondly, we coupled the cardiotoxicity platform with software for predicting hepatotoxicity, immunotoxicity, mutagenicity, reproductive toxicity and tumorigenicity, in order to provide a sufficient filtering method to select safe compounds from any given compound library. A set of 374 artemisinin derivatives, which was retrieved from PubMed (Supplementary Table 1), was used as exemplary case study in the present investigation. Overall, the main aim to use the artemisinin derivatives is to provide novel non-toxic artemisinin derivatives for further drug development. Step-by-step filtering and scheme for the selection of safe compounds are depicted in Fig. 2. The first round of filtering with our cardiotoxicity model delivered 197 compounds (~ 53%)



**Fig. 2** Computational filtering steps to select safe artemisinin derivatives

(Supplementary Table 2) as potentially safe with low cardiotoxicity features. As a next filtering step, mutagenicity, tumorigenicity and reproductive toxicity were assessed by using DataWarrior software. Here, 17 compounds out of 197 were selected (Supplementary Table 3). The last filtering step was performed by using the Pro-Tox II software focusing on hepatotoxicity and immunotoxicity, and finally resulted in 7 out of 17 compounds (deoxydihydroartemisinin, 3-hydroxy-deoxy-dihydroartemisinin, 3-desoxy-dihydroartemisinin, dihydroartemisinin-furano acetate-d3, deoxyartemisinin, artemisinin G, artemisinin B).

A widely accepted marker to evaluate the interaction of chemical compounds with cardiotoxicity is hERG (Bhat and Houghton 2018; Luo et al. 2014).

As a third step in the prediction process, we chose to perform an *in silico* evaluation of the models for cardiotoxicity and the additional parameters (mutagenicity, tumorigenicity, reproductive effect, immunotoxicity and hepatotoxicity) with a list of 39 compounds with known cardiotoxicity taken from the literature (Mladenka et al. 2018). Here, a 97.4% success rate was achieved, as 38 of them were predicted to involve toxicity. Thirty-four compounds were successfully predicted as cardiotoxic (Table 2), and the next step of filtering with the additional parameters led to the prediction of four more compounds as toxic (malathion: mutagenicity, tumorigenicity, reproductive toxicity; methyl dopa: reproductive toxicity; grayanotoxin III 6,14 diacetate: immunotoxicity; sorafenib: immunotoxicity and hepatotoxicity).

These results made us confident to select the remaining set of 7 compounds of artemisinin derivatives from the machine learning-based toxicity prediction platform for further experimental *in vitro* and *in vivo* validation.

### Verification of cardiotoxicity in *in silico* predictions

A widely accepted marker to evaluate the interaction of chemical compounds with cardiotoxicity is hERG (Bhat and

**Table 2** Performance of the *in silico* cardiotoxicity prediction

Name	Predicted cardiotoxicity	Name	Predicted cardiotoxicity
5-Fluorouracil	Yes	Lidocain	Yes
Aconitine	Yes	Malathion*	No
Amifostine	Yes	Metamphetamine	Yes
Amphetamine	Yes	Methyl dopa*	No
Atropine	Yes	Methyltestosterone	Yes
Chloroprocaine	Yes	Metoprolol	Yes
Clenbuterol	Yes	Milrinone	No
Clonidine	Yes	Nortryptiline	Yes
Cocaine	Yes	Paclitaxel	Yes
Digoxin	Yes	Phenylephrine	Yes
Dobutamine	Yes	Prednisone	Yes
Doxazosin	Yes	Reserpine	Yes
Doxorubicin	Yes	Ritodrine	Yes
Grayanotoxin iii 6,14 diacetate*	Yes	Rofecoxib	Yes
Ibuprofen	Yes	Salbutamol	Yes
Ibutilide	Yes	Sildenafil	Yes
Isoprenaline	Yes	Sorafenib*	No
Lapatinib	Yes	Theohylline	Yes
Levosimendan	Yes	Verapamil	Yes
		Yohimbine	Yes

Compounds labeled with \* exert mutagenicity, tumorigenicity, reproductive effect, immunotoxicity or hepatotoxicity

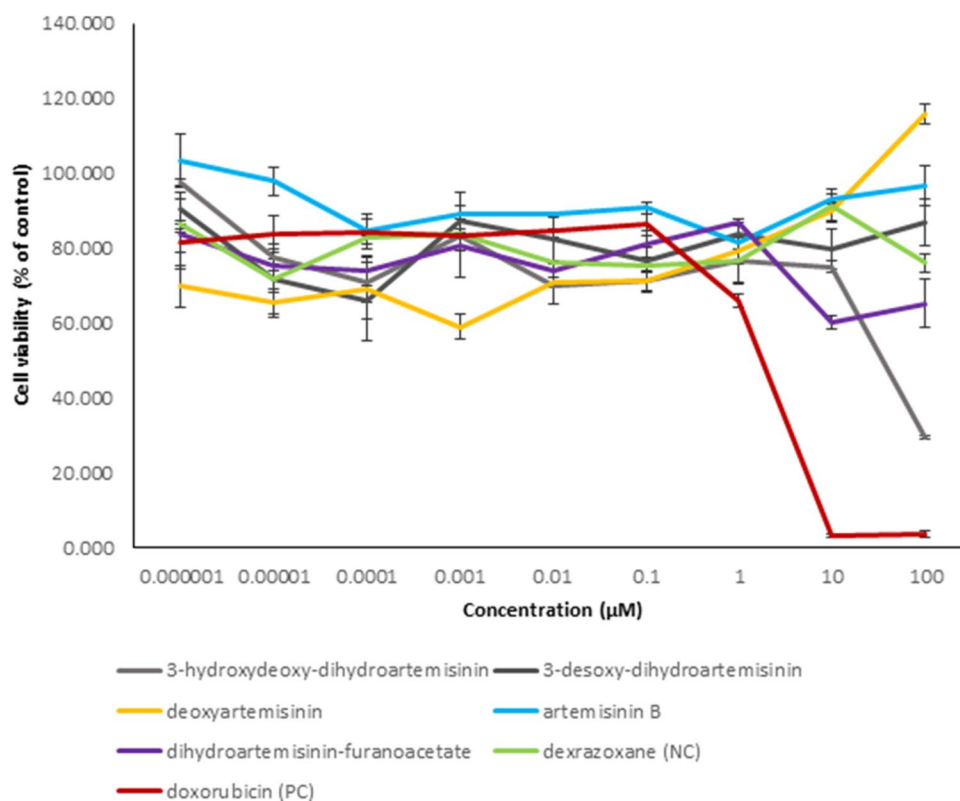
Houghton 2018; Luo et al. 2014). Provided that a compound has a weak interaction with hERG, a safe profile in terms of cardiotoxicity can be assumed (Bhat and Houghton 2018; Lee et al. 2019). As a first step, the interactions of the investigational compounds with hERG were evaluated by means of molecular docking. All compounds except deoxyartemisinin and artemisinin G revealed weaker binding than doxorubicin (Table 3), which is a well-known cardiotoxic drug, also being labeled cardiotoxic in the *in silico* prediction testing (Table 2). Dihydroartemisinin furano acetate-d3 and artemisinin B revealed the weakest interaction with hERG, which can be taken as a clue for missing cardiotoxicity. Remarkably, artemisinin B possessed even weaker binding than the cardioprotective control drug, dexrazoxane (Table 3).

Five out of the 7 selected compounds were available to us to investigate cytotoxicity *in vitro* towards AC16 cardiomyocytes as a parameter to measure cellular cardiotoxicity. As shown in Fig. 3, none of the selected compounds revealed cytotoxicity. The percentages of cell viabilities were quite comparable to those of the cardioprotective control compound, dexrazoxane, except of 3-hydroxydeoxydihydroartemisinin. Here, treatment with a high concentration of 100  $\mu$ M revealed a moderate cytotoxic effect ( $IC_{50}$ :

**Table 3** Molecular docking results of selected artemisinin derivatives on hERG

Compound	LBE (kcal/mol)	Predicted inhibition constant ( $\mu\text{M}$ )
Deoxydihydro-artemisinin	$-5.050 \pm 0.000$	$198.777 \pm 0.235$
3-Hydroxydeoxy-dihydroartemisinin	$-5.053 \pm 0.006$	$197.417 \pm 0.556$
3-Desoxy-dihydroartemisinin	$-5.050 \pm 0.000$	$200.210 \pm 0.113$
Dihydroartemisinin-furanoacetate	$-4.893 \pm 0.006$	$258.980 \pm 1.825$
Deoxyartemisinin	$-5.230 \pm 0.000$	$146.603 \pm 0.391$
Artemisinin G	$-5.200 \pm 0.000$	$155.018 \pm 0.295$
Artemisinin B	$-4.467 \pm 0.031$	$531.677 \pm 29.573$
Doxorubicin (positive control)	$-5.160 \pm 0.066$	$166.230 \pm 17.840$
Dexrazoxane (negative control)	$-4.570 \pm 0.000$	$449.210 \pm 1.417$

$24.915 \pm 0.247 \mu\text{M}$ ). Considerable cytotoxicity towards AC16 cardiomyocytes was solely observed with doxorubicin. Here, the  $\text{IC}_{50}$  value was  $1.918 \pm 0.230 \mu\text{M}$ . It was not possible to measure  $\text{IC}_{50}$  values for all other artemisinin derivatives as well as dexrazoxane in concentrations up to  $100 \mu\text{M}$ . Even cell viability at high concentrations for artemisinin B and deoxyartemisinin was similar to cytotoxicity results previously reported for spinochrome D on AC16 cell line (Yoon et al. 2018). Overall, these results indicate that

**Fig. 3** Cytotoxicity assessment of the selected compounds on AC16 cardiomyocytes. Results represent mean values and standard deviation of each three independent experiments with each 6 parallel measurements

the six artemisinin derivatives may reveal a sufficient cardiotoxicity profile for their use in vivo.

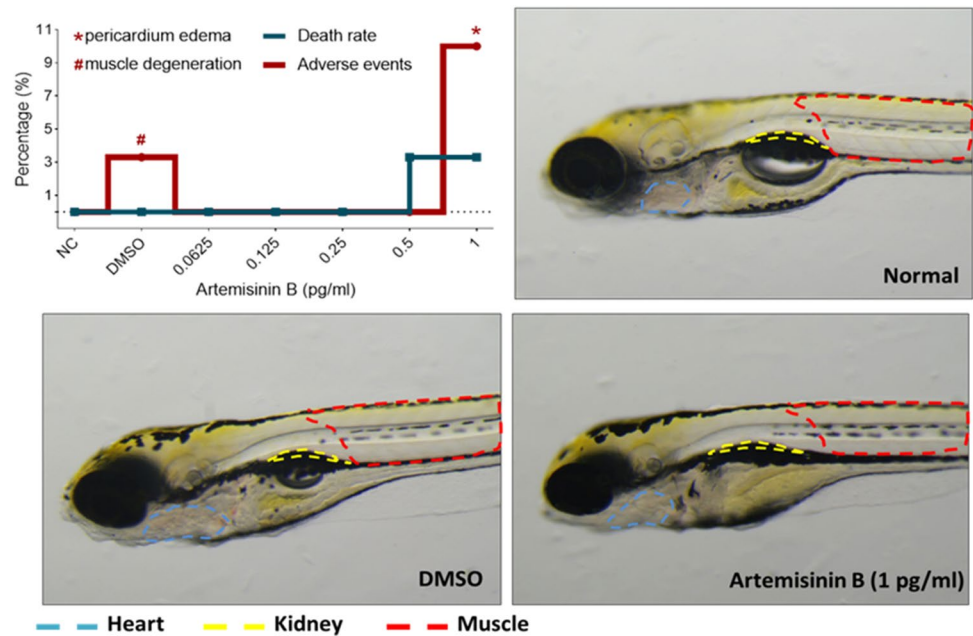
### Toxicity evaluation of artemisinin B

The compound artemisinin B with the best performance in the molecular docking and cell viability experiments was finally tested in the Zebrafish model. The curves of death rate and adverse events of zebrafish larvae are shown in Fig. 4. Zebrafish death was observed with artemisinin B at  $0.5 \text{ pg/ml}$ , and only one dead fish (3.3% death rate) was found. A death rate below 10% can be regarded as experimental error. Pericardial edema was observed as the only possible adverse event induced by artemisinin B at  $1.0 \text{ pg/ml}$ . Since the incidence rate of adverse event was less than 10% which can be accepted as normal response,  $1.0 \text{ pg/ml}$  can still be estimated as safe dose of artemisinin B. Therefore, artemisinin B at a dose range within  $1.0 \text{ pg/ml}$  was non-toxic to zebrafish larvae.

### Transcriptomic analyses

Transcriptome-wide RNA expression analyses revealed that artemisinin B treatment of AC16 cardiomyocytes led to deregulations in cardiotoxicity marker genes in a similar manner as dexrazoxane, and it caused the upregulation of various cardiotoxicity marker genes, which are known to

**Fig. 4** Mortality and adverse events of zebrafish larvae upon artemisinin B treatment



be associated with doxorubicin associated cardiotoxicity. Referring to the literature, where potential cardiotoxicity marker genes were reported (Yoon et al. 2018) and considering those genes, *APP* was 644-fold upregulated, *CCND1* was 385-fold upregulated, *NT5E* was 774-fold upregulated, *PCNA* was 126-fold upregulated, *PRNP* was 360-fold upregulated, *STK39* was 60-fold upregulated, *TXNIP* was 371-fold upregulated by artemisinin B.

*AIFM1* was 170-fold downregulated, *APP* was 633-fold downregulated, *ARAF* was 675-fold downregulated, *CBR1* was 650-fold downregulated, *CNP* was 160-fold downregulated, *DLD* was 394-fold downregulated, *GSK3B* was 697-fold downregulated, *NBN* was 189-fold downregulated, *NT5E* was 510-fold downregulated, *PBX3* was 222-fold downregulated, *RAB2A* was 788-fold downregulated, *SP110* was 264-fold downregulated, *UBE4B* was 540-fold downregulated, *VIM* was 558-fold downregulated by doxorubicin.

*CTSC* was 291-fold upregulated, *DVLI* was 66-fold upregulated, *KLF6* was 163-fold upregulated, *MUS81* was 50-fold upregulated, *NT5E* was 603-fold upregulated, *PFNI* was 562-fold upregulated, *PRNP* was 283-fold upregulated by dexrazoxane (Table 4).

Seventy-eight genes were commonly deregulated in artemisinin B and dexrazoxane and 67 of them are both down- or upregulated in artemisinin B and dexrazoxane-treated cells (Supplementary Table 4). Twenty-eight genes are commonly deregulated in all treatments and 9 out of those 28 genes are reversely deregulated in artemisinin B and dexrazoxane-treated cells compared to doxorubicin-treated cells (Table 5), implying cardiotoxicity features of these genes (*ANPEP*, *BACE1*, *C1QBP*, *MGLL*, *MT1A*, *MT1B*, *NT5E*, *RGS4*, *RPL23A*).

**Table 4** Fold change of cardiotoxicity marker genes in artemisinin B-, dexrazoxane- and doxorubicin-treated AC16 cells

Gene	Artemisinin B	Gene	Dexrazoxane	Gene	Doxorubicin
<i>APP</i>	644.150	<i>CTSC</i>	291.693	<i>AIFM1</i>	- 170.564
<i>CCND1</i>	385.042	<i>DVLI</i>	66.551	<i>APP</i>	- 633.831
<i>NT5E</i>	774.197	<i>KLF6</i>	163.719	<i>ARAF</i>	- 675.091
<i>PCNA</i>	126.363	<i>MUS81</i>	50.017	<i>CBR1</i>	- 650.829
<i>PRNP</i>	360.351	<i>NT5E</i>	603.894	<i>CNP</i>	- 160.800
<i>STK39</i>	60.882	<i>PFNI</i>	562.100	<i>DLD</i>	- 394.123
<i>TXNIP</i>	371.243	<i>PRNP</i>	283.204	<i>GSK3B</i>	- 697.101
				<i>NBN</i>	- 189.766
				<i>NT5E</i>	- 510.647
				<i>PBX3</i>	- 222.206
				<i>RAB2A</i>	- 788.420
				<i>SP110</i>	- 264.249
				<i>UBE4B</i>	- 540.823
				<i>VIM</i>	- 558.631

## Discussion

Recent advances of in silico approaches and tools are promising in terms of systematic evaluation of drug-induced cardiovascular (CV) complications in both drug discovery and post-marketing surveillance (Collins et al. 2015; Lorberbaum et al. 2016a, 2016b). Recently, a machine learning-based cardiotoxicity prediction platform has been reported (Cai et al. 2018). In the present study, we generated models based on random forest algorithm

**Table 5** Fold change of reversely deregulated genes in artemisinin B- or dextrazoxane-treated AC16 cells compared to doxorubicin-treated AC16 cells.

Gene	Artemisinin B	Dextrazoxane	Doxorubicin
<i>ANPEP</i>	503.888	154.716	– 1166.090
<i>BACE1</i>	– 60.065	– 38.736	131.002
<i>CIQBP</i>	251.246	296.638	– 1196.785
<i>MGLL</i>	41.710	205.947	– 229.316
<i>MT1A</i>	3050.855	2360.951	– 2028.899
<i>MT1B</i>	1895.162	1362.674	– 1498.303
<i>NT5E</i>	774.197	603.894	– 510.647
<i>RGS4</i>	256.294	348.026	– 361.650
<i>RPL23A</i>	– 611.567	– 612.719	676.744

with leave-one-out cross-validation to predict cardiotoxicities of different categories in a similar manner as described by Cai et al. (2018). Moreover, we considered other toxicities as well with the use of additional software applications. Our models outperformed in all categories in learning steps except for specificity (0.779 vs. 0.813) and precision (0.780 vs. 0.794) of the heart block model and specificity (0.759 vs. 0.765) of the myocardial infarction model. Regarding the external validation set, we outperformed all categories except for specificity (0.676 vs. 0.710) of the hypertension model. Seven compounds of artemisinin derivatives (deoxydihydro-artemisinin, 3-hydroxy-deoxy-dihydroartemisinin, 3-desoxy-dihydroartemisinin, dihydroartemisinin-furano acetate-d3, deoxyartemisinin, artemisinin G, artemisinin B) passed the toxicity filtering process not only as non-cardiotoxic compounds but also as safe compounds in terms of hepatotoxicity, mutagenicity, tumorigenicity, and reproductive toxicity by including those additional toxicity parameters. Even if drug candidates exert promising pharmacological features towards their therapeutic disease-specific target, non-tolerable toxicities may lead to a drop out of these compounds from the drug development pipeline. For this purpose, suitable toxicity profiling of candidate drugs is essential. The presented toxicity prediction platform was highly accurate and reliable and may, therefore, offer the opportunity to speed up the preclinical drug discovery process and to reduce the developmental costs. The platform proved its success rate in external validation sets, and selected artemisinin derivatives revealed weak interaction with hERG and a comparable safety profile with that of a cardioprotective agent, dextrazoxane, on cardiomyocytes.

Drug attrition and thus approval failure in the last 20 years were mainly due to the lack of efficacy (accounting for ~ 30% of failures) and safety (toxicology and clinical safety accounting for a further ~ 30% (Ferri et al. 2013)). Cardiovascular toxicity is among the most frequent serious

adverse drug reactions and cause of withdrawal for marketed drugs. Therefore, it is of utmost importance to identify and characterize the cardiotoxicity risk of drug candidates as early as possible at the preclinical stage of development (Redfern et al. 2010). Safety liabilities associated with the cardiovascular system account for 45% of the total post-approval drug withdrawal from the market, compared to 32% for the hepatic system (Stevens and Baker 2009). Such a high incidence and severity of cardiovascular toxicity in the late-stage of clinical development can lead to restrictions in indications and dose levels, pre- and/or post-approval monitoring, and ultimately drug discontinuation or withdrawal (Ferri et al. 2013).

Considering the cardiotoxicity marker genes, which are deregulated in doxorubicin-treated cardiomyocytes (Yoon et al. 2018), artemisinin B treatment caused an upregulation of 7 genes (*APP*, *CCND1*, *NT5E*, *PCNA*, *PRNP*, *STK39*, *TXNIP*), while treatment with the cardioprotective control, dextrazoxane, revealed an upregulations of another 5 genes (*CTSC*, *DVLI*, *KLF6*, *MUS81*, *PFNI*), but interestingly also *NT5E* and *PRNP*. By contrast, doxorubicin led to the downregulation of 14 genes (*AIFM1*, *APP*, *ARAF*, *CBR1*, *CNP*, *DLG*, *GSK3B*, *NBN*, *NT5E*, *PBX3*, *RAB2A*, *SP110*, *UBE4B*, *VIM*).

Nine genes (including *NT5E*) were reversely deregulated in artemisinin B- or dextrazoxane-treated cells compared to doxorubicin-treated cells, implying cardiotoxicity marker properties of these genes (*ANPEP*, *BACE1*, *CIQBP*, *MGLL*, *MT1A*, *MT1B*, *NT5E*, *RGS4*, *RPL23A*) in addition to the genes mentioned above. *RGS4*-null mice showed a reduced basal heart rate (Stewart et al. 2012) and enhanced bradycardia resulting from *RGS4* loss due to intrinsic alterations in cardiac automaticity (Cifelli et al. 2008). Artemisinin B and dextrazoxane treatments led to the upregulation of *RGS4*, whereas doxorubicin downregulated *RGS4*, implying cardioprotection by artemisinin B. Among the previously reported doxorubicin-related downregulated cardiotoxicity marker genes (Todorova et al. 2012), *APP*, *CCND1*, *NT5E*, *PCNA*, *PRNP*, *STK39*, *TXNIP* were upregulated in artemisinin B-treated cardiomyocytes, further supporting the cardioprotective properties of artemisinin B.

The cardiotoxic potential of artemisinin derivatives has been reported in dogs (Yin et al. 2014). However, human trials did not show considerable heart function impairment signs by artemisinin derivatives (Efferth and Kaina 2010). This is presumably due to the low dose of artesunate applied in malarial therapy (Efferth and Kaina 2010). The therapeutically effective doses may be different for the treatment of cancer or other diseases. Therefore, the safety of artemisinins in malaria treatment does not allow to extrapolate on safety for other diseases. Only via dose escalation studies in animal and human studies, the



appropriate cardiotoxic profiles of artemisinin derivatives can be reliably determined.

Radioligand binding assays, electrophysiology measurements, rubidium-flux assays, and fluorescence-based assays are among the methods to assess cardiotoxicity (Polak et al. 2009). Zebrafish modeling is another widely used strategy in that regard (Sarmah and Marrs 2016). However, such methods are less feasible for the evaluation of a large number of compounds in early-stage drug discovery because of high expenses and low throughput. After *in silico* filtering and *in vitro* experimentation, we selected artemisinin B and tested it in a zebrafish model, because artemisinin B revealed the weakest interaction with hERG and the lowest cytotoxicity towards AC16 cardiomyocytes. Our results indicate that artemisinin B at 1.0 pg/ml was non-toxic to zebrafish larvae. We conclude that artemisinin B is a promising non-cardiotoxic compound, since it revealed a similar gene expression profile to that of dexrazoxane and showed only slight cardiotoxicity in the picomolar range in zebrafish.

## Conclusions

In the present study, we established a toxicity prediction platform involving cardiotoxicity, hepatotoxicity, immunotoxicity, tumorigenicity, mutagenicity, and reproductive toxicity. Random forest feature selection algorithm was used for the establishment of cardiotoxicity models, and high AUC scores above 0.830 were achieved for all five cardiotoxicity indications. Using artemisinin derivatives as example, 7 out of 374 compounds passed the toxicity filtering procedure. Selected compounds revealed weak interaction with hERG, low toxicity towards cardiomyocytes, verifying their safe profile. Artemisinin B treatment revealed a similar gene expression profile to that of dexrazoxane and showed only slight cardiotoxicity in the picomolar range in zebrafish. In conclusion, our toxicity prediction scheme can speed up the identification of safe compounds in preclinical drug development.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00204-021-03058-4>.

**Acknowledgements** The authors are grateful to Dr. James Spiers and his PhD student Sylwia Konieczna (Trinity College Dublin, Department of Pharmacology and Therapeutics) for providing the AC16 human cardiomyocyte cell line. We thank the Microarray Unit of the Genomics and Proteomics Core Facility, German Cancer Research Center (DKFZ), for providing excellent expression profiling services. Parts of this research were conducted using the supercomputer MOGON and/or advisory services offered by Johannes Gutenberg University Mainz (hpc.uni-mainz.de), which is a member of the AHRP (Alliance for High-Performance Computing in Rhineland Palatinate, [www.ahrp.info](http://www.ahrp.info)) and the Gauss Alliance e.V. The authors gratefully acknowledge the computing time granted on the supercomputer MOGON II at the Johannes Gutenberg University Mainz (hpc.uni-mainz.de).

**Author contributions** Formal analysis, OK; Investigation, OK; Methodology, OK, LS, EF, SMK; Supervision, TE; Writing—original draft, OK; Writing—review and editing, LS, SMK, TE.

**Funding** Open Access funding enabled and organized by Projekt DEAL. Onat Kadioglu is funded by an intramural fund at the Johannes Gutenberg University Mainz, Germany.

**Availability of data and material** All data generated or analyzed during this study are included in this published article.

## Declaration

**Conflict of interest** The authors declare that there is no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Banerjee P, Eckert AO, Schrey AK, Preissner R (2018) ProTox-II: a webserver for the prediction of toxicity of chemicals. *Nucleic Acids Res* 46(W1):W257–W263. <https://doi.org/10.1093/nar/gky318>
- Bhat R, Houghton M (2018) A human cardiomyocyte cell-line expressing hERG: an improved system for testing drug-associated hERG blocking and cardiotoxicity. *J Pharmacol Tox Met* 93:114–114. <https://doi.org/10.1016/j.vascn.2018.01.391>
- Briggs K, Cases M, Heard DJ et al (2012) Inroads to predict *in vivo* toxicology—an introduction to the eTOX Project. *Int J Mol Sci* 13(3):3820–3846. <https://doi.org/10.3390/ijms13033820>
- Cai C, Fang J, Guo P et al (2018) *In silico* pharmacoepidemiologic evaluation of drug-induced cardiovascular complications using combined classifiers. *J Chem Inf Model* 58(5):943–956. <https://doi.org/10.1021/acs.jcim.7b00641>
- Cheng YJ, Wu R, Cheng ML et al (2017) Carboplatin-induced hematotoxicity among patients with non-small cell lung cancer: Analysis on clinical adverse events and drug-gene interactions. *Oncotarget* 8(19):32228–32236. <https://doi.org/10.18632/oncotarget.12951>
- Cifelli C, Rose RA, Zhang H et al (2008) RGS4 regulates parasympathetic signaling and heart rate control in the sinoatrial node. *Circ Res* 103(5):527–535. <https://doi.org/10.1161/CIRCRESAHA.108.180984>
- Collins TA, Bergenholm L, Abdulla T et al (2015) Modeling and simulation approaches for cardiovascular function and their role in safety assessment. *Cpt-Pharmacomet Syst* 4(3):175–188. <https://doi.org/10.1002/psp4.18>
- Demsar J, Curk T, Erjavec A et al (2013) Orange: data mining toolbox in python. *J Mach Learn Res* 14:2349–2353

- Efferth T (2017a) Cancer combination therapies with artemisinin-type drugs. *Biochem Pharmacol* 139:56–70. <https://doi.org/10.1016/j.bcp.2017.03.019>
- Efferth T (2017b) From ancient herb to modern drug: artemisia annua and artemisinin for cancer therapy. *Semin Cancer Biol* 46:65–83. <https://doi.org/10.1016/j.semcancer.2017.02.009>
- Efferth T, Kaina B (2010) Toxicity of the antimalarial artemisinin and its derivatives. *Crit Rev Toxicol* 40(5):405–421. <https://doi.org/10.3109/10408441003610571>
- Efferth T, Bauer R, Funk JO, Davey M, Volm M, Davey R (2002) Molecular modes of action of antimalarial artemisinin derivatives as novel anticancer drugs. *Eur J Cancer* 38:S99–S99
- Efferth T, Romero MR, Wolf DG, Stamminger T, Marin JGG, Marschall M (2008) The antiviral activities of artemisinin and artesunate. *Clin Infect Dis* 47(6):804–811. <https://doi.org/10.1086/591195>
- Elgenaidi IS, Spiers JP (2019) Hypoxia modulates protein phosphatase 2A through HIF-1 alpha dependent and independent mechanisms in human aortic smooth muscle cells and ventricular cardiomyocytes. *Brit J Pharmacol* 176(11):1745–1763. <https://doi.org/10.1111/bph.14648>
- Ferri N, Siegl P, Corsini A, Herrmann J, Lerman A, Benghozi R (2013) Drug attrition during pre-clinical and clinical development: understanding and managing drug-induced cardiotoxicity. *Pharmacol Ther* 138(3):470–484. <https://doi.org/10.1016/j.pharmthera.2013.03.005>
- Issa NT, Wathieu H, Ojo A, Byers SW, Dakshnamurthy S (2017) Drug metabolism in preclinical drug development: a survey of the discovery process, toxicology, and computational tools. *Curr Drug Metab* 18(6):556–565. <https://doi.org/10.2174/1389200218666170316093301>
- Kadioglu O, Efferth T (2019) A machine learning-based prediction platform for p-glycoprotein modulators and its validation by molecular docking. *Cells*. <https://doi.org/10.3390/cells8101286>
- Lee WM (2003) Medical progress: drug-induced hepatotoxicity. *New Engl J Med* 349(5):474–485. <https://doi.org/10.1056/NEJMr021844>
- Lee HM, Yu MS, Kazmi SR et al (2019) Computational determination of hERG-related cardiotoxicity of drug candidates. *BMC Bioinform*. <https://doi.org/10.1186/s12859-019-2814-5>
- Liesse K, Harris J, Chan M, Schmidt ML, Chiu B (2018) Dexrazoxane significantly reduces anthracycline-induced cardiotoxicity in pediatric solid tumor patients: a systematic review. *J Pediatr Hematol Oncol* 40(6):417–425. <https://doi.org/10.1097/MPH.0000000000001118>
- Lopez-Lopez E, Naveja JJ, Medina-Franco JL (2019) DataWarrior: an evaluation of the open-source drug discovery tool. *Expert Opin Drug Dis* 14(4):335–341. <https://doi.org/10.1080/17460441.2019.1581170>
- Lorberbaum T, Sampson KJ, Chang JB et al (2016a) Coupling data mining and laboratory experiments to discover drug interactions causing QT prolongation. *J Am Coll Cardiol* 68(16):1756–1764. <https://doi.org/10.1016/j.jacc.2016.07.761>
- Lorberbaum T, Sampson KJ, Woosley RL, Kass RS, Tatonetti NP (2016b) An integrative data science pipeline to identify novel drug interactions that prolong the QT interval. *Drug Saf* 39(5):433–441. <https://doi.org/10.1007/s40264-016-0393-1>
- Luo F, Gu JY, Chen LR, Xu XJ (2014) Molecular docking and molecular dynamics studies on the structure-activity relationship of fluoroquinolone for the HERG channel. *Mol Bio Syst* 10(11):2863–2869. <https://doi.org/10.1039/c4mb00396a>
- McGowan JV, Chung R, Maulik A, Piotrowska I, Walker JM, Yellon DM (2017) Anthracycline chemotherapy and cardiotoxicity. *Cardiovasc Drug Ther* 31(1):63–75. <https://doi.org/10.1007/s10557-016-6711-0>
- Michaelsen FW, Saeed MEM, Schwarzkopf J, Efferth T (2015) Activity of Artemisia annua and artemisinin derivatives, in prostate carcinoma. *Phytomed Int J Phytotherapy Phytopharmacol* 22(14):1223–1231. <https://doi.org/10.1016/j.phymed.2015.11.001>
- Mladenka P, Applova L, Patocka J et al (2018) Comprehensive review of cardiovascular toxicity of drugs and related agents. *Med Res Rev* 38(4):1332–1403. <https://doi.org/10.1002/med.21476>
- Morris GM, Huey R, Lindstrom W et al (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem* 30(16):2785–2791. <https://doi.org/10.1002/jcc.21256>
- Mu XZ, Wang CC (2018) Artemisinins—a promising new treatment for systemic lupus erythematosus: a descriptive review. *Curr Rheumatol Rep*. <https://doi.org/10.1007/s11926-018-0764-y>
- Narayan HK, Putt ME, Kosaraju N et al (2019) Dexrazoxane preferentially mitigates doxorubicin cardiotoxicity in female children with sarcoma. *Open Heart* 6(1):e001025. <https://doi.org/10.1136/openhrt-2019-001025>
- Newman DJ, Cragg GM (2007) Natural products as sources of new drugs over the last 25 years. *J Nat Prod* 70(3):461–477. <https://doi.org/10.1021/np068054v>
- Oliveira PA, Colaco A, Chaves R, Guedes-Pinto H, De-La-Cruz LF, Lopes C (2007) Chemical carcinogenesis. *An Acad Bras Cienc* 79(4):593–616. <https://doi.org/10.1590/S0001-3765200700400004>
- Ooko E, Saeed MEM, Kadioglu O et al (2015) Artemisinin derivatives induce iron-dependent cell death (ferroptosis) in tumor cells. *Phytomed Int J Phytother Phytopharmacol* 22(11):1045–1054. <https://doi.org/10.1016/j.phymed.2015.08.002>
- Ooko E, Alsalam T, Saeed B et al (2016) Modulation of P-glycoprotein activity by novel synthetic curcumin derivatives in sensitive and multidrug-resistant T-cell acute lymphoblastic leukemia cell lines. *Toxicol Appl Pharm* 305:216–233. <https://doi.org/10.1016/j.taap.2016.06.002>
- Polak S, Wisniewska B, Brandys J (2009) Collation, assessment and analysis of literature in vitro data on hERG receptor blocking potency for subsequent modeling of drugs' cardiotoxic properties. *J Appl Toxicol* 29(3):183–206. <https://doi.org/10.1002/jat.1395>
- Redfern W, Ewart L, Hammond T et al (2010) Impact and frequency of different toxicities throughout the pharmaceutical life cycle. *The Toxicologist* 114(S1):1081
- Sander T, Freyss J, von Korff M, Rufener C (2015) Data warrior: an open-source program for chemistry aware data visualization and analysis. *J Chem Inf Model* 55(2):460–473. <https://doi.org/10.1021/ci500588j>
- Sarmah S, Marrs JA (2016) Zebrafish as a vertebrate model system to evaluate effects of environmental toxicants on cardiac development and function. *Int J Mol Sci*. <https://doi.org/10.3390/ijms17122123>
- Shi Z, Chen Y, Lu C et al (2018) Resolving neuroinflammation, the therapeutic potential of the anti-malaria drug family of artemisinin. *Pharmacol Res* 136:172–180. <https://doi.org/10.1016/j.phrs.2018.09.002>
- Stevens JL, Baker TK (2009) The future of drug safety testing: expanding the view and narrowing the focus. *Drug Discov Today* 14(3–4):162–167. <https://doi.org/10.1016/j.drudis.2008.11.009>
- Stewart A, Huang J, Fisher RA (2012) RGS proteins in heart: brakes on the vagus. *Front Physiol* 3:95. <https://doi.org/10.3389/fphys.2012.00095>
- Todorova VK, Beggs ML, Delongchamp RR et al (2012) Transcriptome profiling of peripheral blood cells identifies potential

- biomarkers for doxorubicin cardiotoxicity in a rat model. *PLoS ONE* 7(11):e48398. <https://doi.org/10.1371/journal.pone.0048398>
- Wang W, MacKinnon R (2017) Cryo-EM structure of the open human ether-a-go-go-Related K(+) channel hERG. *Cell* 169(3):422–430 e10 <https://doi.org/10.1016/j.cell.2017.03.048>
- Yan G, Dawood M, Bockers M et al (2020) Multiple modes of cell death in neuroendocrine tumors induced by artesunate. *Phytomedicine* 79:153332. <https://doi.org/10.1016/j.phymed.2020.153332>
- Yin JY, Wang HM, Wang QJ et al (2014) Subchronic toxicological study of two artemisinin derivatives in dogs. *PLoS ONE* 9(4):e94034. <https://doi.org/10.1371/journal.pone.0094034>
- Yoon CS, Kim HK, Mishchenko NP, et al. (2018) Spinochrome D attenuates doxorubicin-induced cardiomyocyte death via improving glutathione metabolism and attenuating oxidative stress. *Mar Drugs*. <https://doi.org/10.3390/md17010002>
- Zhang S, Meng T, Liu J, Zhang X, Zhang J (2015) Cardiac protective effects of dexrazoxane on animal cardiotoxicity model induced by anthracycline combined with trastuzumab is associated with upregulation of calpain-2. *Med (Baltim)* 94(4):e445. <https://doi.org/10.1097/MD.0000000000000445>
- Zheng C, Shan L, Tong P, Efferth T (2020) Cardiotoxicity and cardioprotection by artesunate in larval zebrafish. *Dose Response* 18(1):1559325819897180. <https://doi.org/10.1177/1559325819897180>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.