

Insights into Drug Cardiotoxicity from Biological and Chemical Data: The First Public Classifiers for FDA DICTrank

Srijit Seal^{*1}, *Ola Spjuth*², *Layla Hosseini-Gerami*³, *Miguel García-Ortegón*⁴, *Shantanu Singh*¹,
*Andreas Bender*⁴, *Anne E. Carpenter*¹

¹Imaging Platform, Broad Institute of MIT and Harvard, US

²Department of Pharmaceutical Biosciences, Uppsala University, Sweden

³Ignota Labs, UK

⁴Department of Chemistry, University of Cambridge, UK

Documentation: https://broad.io/DICTrank_Predictor

Code: <https://github.com/srijitseal/DICTrank>

Datasets: 10.6084/m9.figshare.24312274

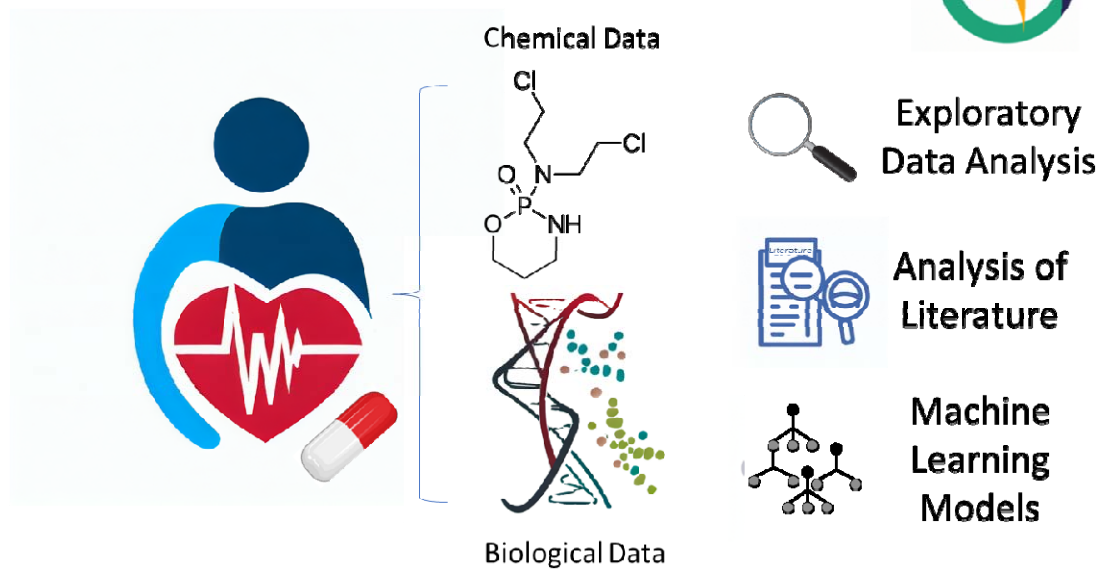
*Correspondence:

seal@broadinstitute.org

Abstract

Drug-induced cardiotoxicity (DICT) is a major concern in drug development, accounting for 10-14% of postmarket withdrawals. In this study, we explored the capabilities of various chemical and biological data to predict cardiotoxicity, using the recently released Drug-Induced Cardiotoxicity Rank (DICTrank) dataset from the United States FDA. We analyzed a diverse set of data sources, including physicochemical properties, annotated mechanisms of action (MOA), Cell Painting, Gene Expression, and more, to identify indications of cardiotoxicity. We found that such data, including protein targets, especially those related to ion channels (such as hERG), physicochemical properties (such as electrotopological state) as well as peak concentration in plasma offer strong predictive ability as well as valuable insights into DICT. We also found compounds annotated with particular mechanisms of action, such as cyclooxygenase inhibition, could distinguish between most-concern and no-concern DICT compounds. Cell Painting features related to ER stress discern the most-concern cardiotoxic compounds from non-toxic compounds. While models based on physicochemical properties currently provide substantial predictive accuracy (AUCPR = 0.93), this study also underscores the potential benefits of incorporating more comprehensive biological data in future DICT predictive models. With the availability of -omics data in the future, using biological data promises enhanced predictability and delivers deeper mechanistic insights, paving the way for safer therapeutic drug development. All models and data used in this study are publicly released at https://broad.io/DICTrank_Predictor

DICTrank Predictor



Introduction

Drug-induced cardiotoxicity (DICT) is a leading cause of drug withdrawals during post-market surveillance. One study showed that 10% of withdrawals in the last 4 decades were due to cardiovascular safety concerns, including previously successful therapeutics such as rofecoxib, tegaserod, sibutramine, and rosiglitazone.¹ Another study found that cardiotoxicity was the third most common reason for adverse drug reactions and accounted for 14% of withdrawals.² Worryingly, the rate of DICT-related withdrawals may even be increasing, accounting for 17 out of 38 cases among drugs approved between 1994 and 2006.^{1,3}

DICT is associated with both functional damage such as arrhythmia, which alters mechanical function, and structural damage such as morphological damage in cardiomyocytes; functional damage and structural damage in the heart can be interrelated, where one may precipitate the other.⁴ DICT can be attributed to several underlying mechanisms affecting myocardial functions and viabilities.⁵ Some drugs, such as anthracyclines, inflict direct myocyte injury via reactive oxygen species production and compromising DNA replication.⁶ Electrophysiological disruptions, for example, measured in the hERG potassium channel blockers, can lead to arrhythmias by causing QT interval prolongation.⁷ Cardiac energy demands can be affected by drugs that interfere with mitochondrial functionality.¹ Drugs may also adversely influence vascular supply, inducing ischemic conditions.⁸ Intracellular calcium regulation for cardiomyocyte activity can also disrupt its homeostasis, resulting in contractile and rhythm abnormalities.⁹ Furthermore, alterations in growth factors and cytokine balances can induce cardiac conditions like fibrosis, and immunologic drug reactions can also cause cardiotoxicity.^{10,11} Several neurohormonal pathways also offer indirect routes for drug-induced cardiac stress.¹² Notably, a single drug might induce cardiotoxicity via multiple mechanisms, and individual patients' responses (which can often manifest as side effects) can be modulated by genetics, concurrent health conditions, and other medications.¹³

To move beyond a limited focus on specific adverse reactions or related proxy assays for cardiotoxicity, the FDA recently released the Drug-Induced Cardiotoxicity Rank (DICTrank) that categorizes drugs based on their risk of causing cardiotoxicity.¹⁴ Similar to the DILLrank data for liver injury¹⁵, the DICTrank system uses FDA drug labeling to comprehensively categorize 1,318 human drugs into four DICT Concern categories based on their potential risk for cardiotoxicity: (1) Most-DICT-Concern, (2) Less-DICT-Concern, (3) No-DICT-Concern and (4) Ambiguous-DICT-Concern. The DICTrank dataset was generated with an expertise review from the FDA, keyword searches, and manual curation of FDA labeling documents as well as data from clinical trials, post-marketing, and literature surveys.

Predictive models for drug-induced cardiotoxicity (DICT) could save considerable time, resources, and human suffering, with the ultimate goal of preventing adverse events in clinical trials and the post-market stage. However, predicting any *in vivo* effect is not a trivial classification task, and most predictive models are built on proxy endpoints (which are often reduced to binary endpoints) without taking into account *in vivo* parameters such as pharmacokinetic parameters.¹⁶ While no models for DICTrank have been publicly available yet to the best of our knowledge, various studies have predicted proxy *in vitro* assays or side effect data from SIDER (Side Effects Resource), some of which are related to cardiotoxicity.¹⁷ Studies focusing on side effects and proxy targets (such as hERG) are reasonable given that compounds that have cardiac-related indications are more likely to show related side effects as well or activity on ion channels.⁷

Previously it was shown that adverse events data and biological data can be used for identifying mechanism hypotheses leading to cardiotoxicity.¹⁸ Wang et al used LINCS L1000 gene expression features to predict a wide range of drug-induced adverse events from the SIDER dataset.¹⁹ Particularly for acute myocardial infarction, models developed achieved an AUC-ROC of 0.84 when using chemical structural data and 0.76 when using Gene Ontology annotations (compared to 0.5 for random models). Gelano et al used a matrix decomposition

algorithm to predict side effect frequencies for drugs and provide biologically interpretable insights.²⁰ MoleculeNet predictions for SIDER side effects, trained on chemical structure data, range from 0.65 to 0.70 AUC-ROC when using a bypass network, a modified version of a multi-task network.²¹

Most predictive models above were built on chemical structure data as input features. Although certain structural motifs or patterns in a molecule can be indicative of toxic properties and analyzing the chemical structure can flag potential cardiotoxic compounds, such models are often limited in their applicability domain, that is, their accuracy is limited to the chemical space of the training data, and they fail to generalize to markedly different chemical structures. Novel chemical and biological data have been previously used to evaluate side effects in general from the SIDER dataset.²² Previous studies have shown that Random Forest models trained on a combination of biological, chemical, and phenotypic features achieved an AUCPR of 0.76 for cardiac disorders.²³

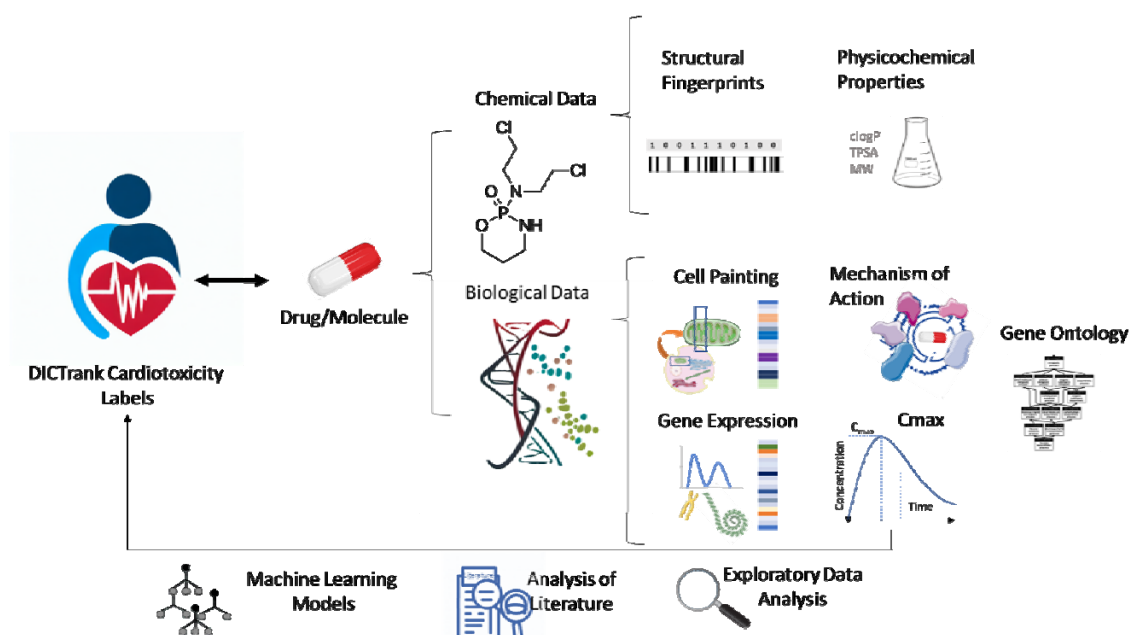


Figure 1: Chemical and biological data sources were used in this study to perform exploratory data analysis on DICTrank and training predictive models (further details in Table 2).

With the availability of the new DICTrank dataset, we used a novel multi-faceted approach using both chemical and biological data (that considers a multitude of possible mechanisms

that can lead to DICT) intending to better understand and make mechanistic insights into a drug's cardiac safety profile. We evaluated a wide range of chemical and biological information, as shown in Figure 1, to determine which feature space is most predictive of DICTrank and evaluated these feature spaces to build the first predictive models of DICTrank using machine learning. Biological data sources included Cell Painting, gene expression, and Gene Ontology,^{24–28} as well as bioactivity, and annotated mechanisms of action (MOA)²⁹ and pharmacokinetic parameters for the peak unbound and total concentration of a drug molecule in plasma³⁰; these offer an alternate feature space to chemical space.³¹ We aimed to glean insights from which chemical and biological data best capture the carefully curated manual annotations in the DICTrank data. Incorporating data from all these sources as feature spaces for predictive models allows for a multifaceted assessment of a drug's potential cardiotoxicity, potentially enhancing the model's accuracy and reliability. Overall, the use of biological data sources along with chemical data improved detection and offered mechanistic insights into the cardiotoxicity of compounds. The models based on chemical structures and physicochemical characteristics are readily accessible for direct use on <> (owing to the constrained availability of public data for other feature types). All code and data for all models can be found on GitHub (<https://github.com/srijitseal/DICTrank>) for local implementation with further details on https://broad.io/DICTrank_Predictor.

Methods

Data sources

We obtained the DICTrank dataset, as released by Qu et al. which includes comprehensive DICTConcern categories for a diverse set of over 1300 drugs.¹⁴ The SIDER database, a pharmacovigilance resource, contained associations for drugs with side effects.^{21,32} We used data from cardiac disorders from the SIDER dataset to compare concordance with DICTrank and enrich the dataset as described later. To gain insights into the mechanisms of action (MOA) of various drugs, we assessed relevant data from the Drug Repurposing Hub³³ which contained information on 6777 drugs for 1130 MOAs and 2183 known targets. To explore the potential targets of drugs, we incorporated the CellScape target predictions on inhibition/antagonism for 2,094 targets at four concentrations (0.1, 1, 10, and 100 μ M).³⁴ We used morphological profiles from the Cell Painting assay²⁴ which considers the impact of drugs on cellular morphology and function. This dataset contained a range of circa 1700 morphological features for over 15,000 compound perturbations. We obtained gene expression data from LINCS L1000 data which contains over 19,000 drugs as described in Wang et al.¹⁹ This study utilized gene expression features derived from LINCS L1000²⁵ transcriptomic data, capturing changes in 978 landmark genes across diverse human cell lines in response to compound perturbations. Gene Ontology-transformed expression features²⁶, which encode biological processes involved with gene expressions affected by the compound perturbations, were extracted from a dataset containing 4,438 annotated features linked to these compounds in the study.¹⁹ The analysis by Wang et al prioritized the strongest signatures across cell line, concentration, and time point for each compound using Characteristic Direction (CD) and evaluated the enrichment across various gene set libraries via Principal Angle Enrichment Analysis (PAEA).³⁵ Finally we used pharmacokinetic data, specifically the maximum unbound and total concentrations (C_{max}) of 758 drugs in the bloodstream, as compiled by Smith et al.³⁰ This dataset contains C_{max} (unbound) for 534 compounds and C_{max} (total) for 749 compounds.

Standardization of the SMILES

For each dataset, we standardized chemical SMILES iteratively using RDKit³⁶ and MolVS³⁷ functionalities. This includes steps for InChI transformation, molecular cleanup, charge neutralization, tautomer normalization, and final standardization. We carried out up to five iterations of the standardization until a standardized SMILES was finalized, otherwise, we chose the most common SMILES from the counter. Finally, the molecule was protonated at pH 7.4 using DimorphiteDL to reflect its likely state at physiological pH.³⁸ Hence, we obtained a standardized SMILES and a standardized InChI.

Preprocessing data

For the DICTrank dataset, we binarized the dataset considering DICT no-concern as 0 and less- and most-concern as 1 as DICTrank labels for machine learning classifiers. We removed compounds that were ambiguous and treated a compound as toxic if there was at least one record of toxicity among duplicates. For the SIDER dataset, we removed duplicate standardized smiles, and similar to the above, labeled a compound as toxic if there was at least one evidence of toxicity among the duplicates. Labels from both SIDER and DICTrank are described in Table 1.

Table 1: Distribution of compound toxicity labels related to cardiotoxicity/ cardiac disorders for all unique compounds from each of the datasets used in this study.

Dataset	Label	Number of Toxic Compounds	Number of Non-Toxic Compounds	Description
SIDER	Cardiac disorders (binary)	829	360 (Absence of Evidence)	Recorded adverse drug reactions from marketed medicines.
DICTrank	DICT Concern Category (Categorical)	Most: 299, Less: 443,	No: 278 (Evidence of Absence)	A ranking system from DICTrank that categorizes drugs according to risk for cardiotoxicity.
	DICTrank label (binary)	742	278 (Evidence of Absence)	Binarized labels obtained from DICT Concern categories used in this study

For the Cell Painting, Gene Expression, and Gene Ontology datasets, we use median cell profiles over standardized SMILES obtaining two datasets: 1783 Cell Painting features for 15,406 compounds, and Gene Expression features for 978 landmark genes and 4,428 Gene Ontology annotations for 9132 compounds. For the MOA dataset, we used one hot encoding of given annotations for compounds, which effectively gives us data for evidence of the presence of MOA/known targets and the absence of evidence. We used a variance threshold of 0.001 to identify and remove low-variance features reducing the dimensionality to 264 MOA and 551 known target features with significant variability. All datasets are released publicly at figshare (10.6084/m9.figshare.24312274) and https://broad.io/DICTrank_Predictor.

Analyzing chemical space overlap between SIDER and DICTrank

We used standardized InChI to calculate the overlap between SIDER and DICTrank datasets. We assessed the physicochemical space using a t-distributed stochastic neighbor embedding (TSNE; as implemented in scikit-learn³⁹) for six physicochemical properties, namely, molecular weight, topological polar surface area, number of rotatable bonds, hydrogen bond donors and acceptor, and the computed logarithm of the partition coefficient. To analyze the chemical space we used a Principal Component analysis (PCA) of the FragFP fingerprints from DataWarrior⁴⁰, which in our experience works better with a higher explained variance in the plot of the principal component analysis compared to Morgan fingerprints.

Structural and physicochemical features

For structural features, we used 2048-bit Morgan Fingerprints as implemented in RDKit.³⁶ For chemical compounds, we computed 1579 descriptors using Mordred.⁴¹ These physicochemical descriptors are derived from 2D representations of compounds, that is, we did not consider 3D descriptors. We removed the descriptors that failed to compute and finally obtained 1038 2-D physicochemical descriptors, and these were used for the machine learning models. For the analysis of feature distributions, we used the full set of 208 RDKit

descriptors, (which are better interpretable compared to Mordred descriptors) as defined in the Descriptors module.³⁶

Predicted targets from CellScape

To derive predicted molecular targets for compounds, we utilized the commercially available CellScape target prediction package (Ignota Labs, 2023).³⁴ This package applies models trained on a mixture of publicly available and proprietary bioactivity data (primarily inhibitory/antagonistic mechanisms) at 0.1, 1, 10, and 100uM with chemical structural features to output a probability score (between 0 and 1) of predicted activity for 2,094 distinct human targets. Although not used in this study, publicly available target prediction alternatives are also available such as PIDGINv4^{42,43} and swisstargetpred⁴⁴. We provide the computed CellScape features for compounds in the DICTrank dataset publicly via figshare (10.6084/m9.figshare.24312274) and https://broad.io/DICTrank_Predictor.

Substructure analysis and retrospective analysis of DrugBank

For substructure analysis, we used SARpy⁴⁵ on the DICTrank dataset, in a method similar to the one applied by Hemmrich et al.⁴⁶. SARpy uses a recursive algorithm for fragmentation. We used two distinct settings for analysis: (1) using both toxic and non-toxic compounds and (2) using only toxic compounds to yield the desired substructures. For both settings, we confined the fragment size within a range of two to 18 atoms, with a minimum occurrence of five times. Furthermore, the positive predictive value (PPV) was adjusted to minimize false negatives. We combined structural alerts from both settings and quantified the frequency of these fragments within the entirety of the DICTrank dataset. We eliminated fragments with a PPV below 0.5. We then manually assessed the remaining fragments, for example removing those having four or fewer atoms, removing substructures like benzene, etc. to obtain 58 structural alerts.

We analyzed all compounds in DrugBank⁴⁷ for the presence of structural alerts from the above to evaluate the risk of the chemical space of drugs for cardiotoxicity. We only used the compounds that did not overlap with the DICTrank dataset for this retrospective analysis to avoid information leaks. We annotated these compounds with labels for cardiac disorders from SIDER and disease area labels from the MOA dataset. We then checked for the presence of structural alerts among the subset of compounds that are currently approved, investigational, experimental, and/or withdrawn drugs.

Analysis of chemical and biological data for differences in feature distribution for DICTrank compounds

We detected features that are predictive of highly cardiotoxic compounds. In order to do this, we detected features for each chemical and biological dataset that had a significant difference in the distribution for the DICT concern categories. For categorical features (SIDER, MOA annotations, and some of the 208 RDKit descriptors), we employed the chi-squared test (as implemented in SciPy⁴⁸) to evaluate the association between categorical variables. We used a contingency table delineating the frequency distribution for each combination of category values. The chi-squared test yielded a statistical value alongside a corresponding p-value. For continuous features (as in the Cell Painting, Gene Expression, Gene Ontology datasets, and some of the 208 RDKit descriptors), we chose the Kruskal-Wallis test (as implemented in SciPy⁴⁸) for evaluating the DICT-Concern labels since it is suited for comparisons involving three or more independent groups. Conversely, when comparing two classes, pairwise, the Mann-Whitney U test (as implemented in SciPy⁴⁸) was used which is adept at discerning differences in distributions between two independent samples. Both tests yield a statistic value alongside its corresponding p-value. For both total unbound/plasma concentrations, as in the Cmax dataset, we used the Mann-Whitney U test to compare the distribution of Cmax among each DICT concern class and the DICTrank label.

Enriching DICTrank compounds with SIDER compounds

We next determined the overlap of compounds (and the concordance in their labels) in the DICTrank dataset with the compounds in SIDER labeled with “Cardiac disorders” using the standardized InChI yielding 776 compounds in common. We next enriched DICTrank with SIDER giving a preference to the DICTrank label in case of a conflict. In this manner, we obtained three datasets besides the DICTrank dataset with the distribution of toxic/non-toxic compounds given in Supplementary Table S1. These are (1) DICTrank, (2) DICTrank enriched with cardiotoxic compounds from SIDER, (3) DICTrank enriched with non-cardiotoxic compounds from SIDER and (4) DICTrank enriched with all compounds from SIDER.

Training predictive models for DICTrank

We trained eleven Random Forest models, each using the following features (as listed in Table 2): (1) Structural fingerprints, (2) Mordred descriptors, (3) MOA labels, (4) MOA labels along with total Cmax, (5) MOA labels along with unbound Cmax, (6) CellScape predicted protein targets, (7) CellScape predicted protein targets along with total Cmax, (8) CellScape predicted protein targets along with unbound Cmax, (9) Cell Painting features, (10) Gene Expression features, and (11) Gene Ontology features.

The training data available for these models depended on the number of compounds for which data was available and varied as given in Supplementary Table S1. As the external test set, we aimed to keep that fixed for a fair evaluation depending on available data as shown in Supplementary Table S2. For models not using Cmax data (where overlaps were larger and hence more data was available), we randomly selected 90 compounds (8.8% of the dataset, 65 cardiotoxic and 21 non-toxic) for which all annotations of feature spaces were available (as described in Supplementary Table S1). These 90 compounds struck a similar balance of DICT concern categories (most: 39, less: 26, and no: 25) as the original

DICTrank dataset. For models using total Cmax data, we used the same external test set comprising 90 compounds since total Cmax data was available for these compounds. However, for models using unbound Cmax data (which had smaller overlaps compared to the above), we used a subset of 78 compounds (57 cardiotoxic and 21 non-toxic) as the external test set as shown in Supplementary Table S2.

Among the models that relied on -omics data (Cell Painting, Gene Expression, and Gene Ontology) we checked for each training compound whether a profile (feature set) was available. If there was no profile available in the respective datasets, we calculated the median profile of all compounds in the original dataset using a v -NN approach, which is different from a fixed k -nn approach; v -nn selects the neighbors based on a condition for each query compound. We used the median profile on the v training compounds that had a Tanimoto similarity greater than 0.70. We ignored any similar compound that appears in the external test set to avoid information leaks. Subsequently, we further discarded any compounds for which no feature profile was found directly or using the above v -nn approach. Thus, while the test sets for the DICTrank and DICTrank enriched datasets are the same, it is important to note that the training data for them vary for the models (as described in Supplementary Table S1) since we dropped compounds where no feature data could be found or matched.

For each of the eleven models, we used a Random Forest classifier, with hyperparameter optimization on the training data using a halving random search with a 5-fold stratified cross-validation with a random oversampling to account for class imbalance (as implemented in scikit-learn³⁹). We used the best hyperparameter-optimized estimator and obtained out-of-fold predictions with a 5-fold stratified cross-validation. We used the out-of-fold predictions and the true labels to optimize the decision threshold for binary classification using the J statistic, calculated as the difference between the true positive rate and the false positive rate. This determines the threshold from ROC curve values where the J statistic is

maximized. The model was finally refitted on the entire training dataset, and we used the optimized threshold to make final predictions based on the predicted probabilities of the external test set.

We trained two ensemble models to combine the models from the eleven feature spaces above. These were based on soft voting, which considered the mean of the scaled predicted probabilities of each mode (scaled according to the best threshold of each model). The first model considered only the six best-performing models (structural, physicochemical, MOA, CellScape, MOA with Cmax total, and CellScape with Cmax total) in the cross-validation (AUC>0.65). The second ensemble model considers all eleven models and thus is evaluated on the reduced external test set of 78 compounds where data from all feature spaces were available.

Model evaluation and applicability domain

We evaluated the classifiers using the balanced accuracy, sensitivity (or recall), specificity, F1 score, Matthews Correlation Coefficient (MCC), AUC-ROC, and the AUC-PR, or precision-recall curve, which focuses on the positive class.

To evaluate the applicability domain of the models, for each compound in the external test set, we calculated the Tanimoto similarity of the nearest neighbor of the same DICTrank label (toxic/non-toxic) in the training dataset. We grouped compounds in 5 equal bins from Tanimoto similarity of 0.0 to 1.0 and evaluated the balanced accuracy and AUCPR in this range for the models used in this study.

Statistics and Reproducibility

We have released the datasets used in this study which are publicly available at [10.6084/m9.figshare.24312274](https://doi.org/10.6084/m9.figshare.24312274). We released the Python code for the models which are publicly available at <https://github.com/srijitseal/DICTrank>.

Results and Discussion

In this study, we used various biological and chemical datasets to discern among the DICT concern categories, driving insights into the carefully annotated FDA DICTrank dataset. We also trained predictive models using these feature spaces. In particular, we used the Cell Painting data from Bray et al, which captures a wide array of cellular phenotypes after perturbation e.g. drug treatment, and has been shown to have a signal for various *in vitro* toxicity.^{24,49} We also used experimental (from the Repurposing Hub²⁹) and predicted bioactivity data derived from models trained on a mixture of publicly available and proprietary datasets (Ignota Labs CellScope³⁴), mostly relating to inhibitory/antagonist mechanisms). For structure-derived feature spaces, we used Morgan fingerprints derived from chemical structures as well as physicochemical Mordred descriptors which are often related to pharmacokinetic properties (such as logD, molecular weight, solubility, permeability, etc.) and implicitly encode the bias between bioactivity classes and chemical structures.⁵⁰ Finally, we looked at pharmacokinetic parameters for the peak unbound and total concentration of a drug molecule in plasma (Cmax).⁵¹ We organized and standardized various chemical and biological data, as shown in Table 2, to analyze their ability to predict DICTrank labels.

Table 2: The description of various feature spaces used in this study.

Feature Space	Dimensions after feature selection (where applicable)	Description	Signal Expected	Source
Chemical Structure	2048-bit vector	ECFP4 (Morgan) fingerprints representing chemical structures	Distinctive patterns of chemical bonding and arrangement	52
Physicochemical properties (Mordred Descriptors)	1038 2-D descriptors	Properties such as lipophilicity, solubility, molecular weight, ionizing potential, etc.	Properties that are associated with negative impacts on ion channels in the heart	41
MOA dataset	264 binary encoded MOAs + 551 known targets	Annotations for mechanism of action and known targets based on knowledge.	Mechanism of action for drugs that inhibit certain ion channels	29
CellScape Target Prediction dataset	1893 predictions for 817 unique targets and concentration combinations (0.1, 1, 10, and 100uM)	Predicted protein target for inhibition/antagonism; does not consider the functionality; prediction is based on chemical structure; updated algorithm from PIDGINv4 ⁴³	Understanding how a drug interacts with various biological targets (not just its primary target) can provide insights into potential off-target effects	34,43
Cell Painting	1783 features	Morphological changes in U2OS cells by a chemical perturbation, using a 5-channel fluorescence microscopy assay	Morphological changes in cells that reflect basic biological processes	24,53
Gene Expression	978 features	Transcriptomic changes in response to chemicals using the L1000 assay	Upregulation or downregulation of genes associated with cardiac stress, apoptosis in cardiac cells, or ion channel function	19,25
Gene Ontology	4438 annotations	Gene Ontology manual annotations based on collective knowledge	Understanding the biological processes, cellular components, and molecular functions affected, e.g., related to cardiac function, cardiac muscle tissue development, or ion homeostasis	19,26
Cmax	2 features	The maximum total and unbound concentration of a drug in plasma	High Cmax would indicate a high risk of cardiotoxicity	30

DICTrank labels are highly concordant with SIDER labels

Among the 776 compounds present in both DICTrank and SIDER cardiac disorders datasets (Figure 2a), we found an 87.24% concordance rate in the annotations (labels) between the two datasets (Supplementary Table S3; SIDER labels have an F1 score of 0.91 when compared against DICTrank labels). This suggests that SIDER labels which ascertain cardiac disorder events reported as associated with each drug and are often dependent on aggregated dispersed public information and package inserts, agree with DICTrank labels which ascertain if a compound is classified as cardiotoxic by the FDA.

The physicochemical space of SIDER and DICTrank generally overlap (Figure 2b), defined as a TSNE space for six physicochemical properties, namely, molecular weight, topological polar surface area, number of rotatable bonds, hydrogen bond donors and acceptor, and the computed logarithm of the partition coefficient. Still, compounds exclusively available in the SIDER dataset could help enrich nontoxic compounds in areas of the chemical space where DICTrank only covers toxic compounds. We see a similar trend for a chemical space defined in fragment fingerprints space from DataWarrior⁴⁰ (Supplementary Figure S1). Therefore, we chose to assess whether adding SIDER compounds to DICTrank compounds improved predictive ability. Interestingly, other categories of SIDER adverse effects were highly correlated to DICTrank (Figure 2c); the interrelationships of vascular disorders and nervous system disorders are well known.^{5,54} Overall, drug adverse events, as recorded in SIDER, have a high concordance with DICTrank labels from the FDA and there is a strong rationale to rely on both resources.

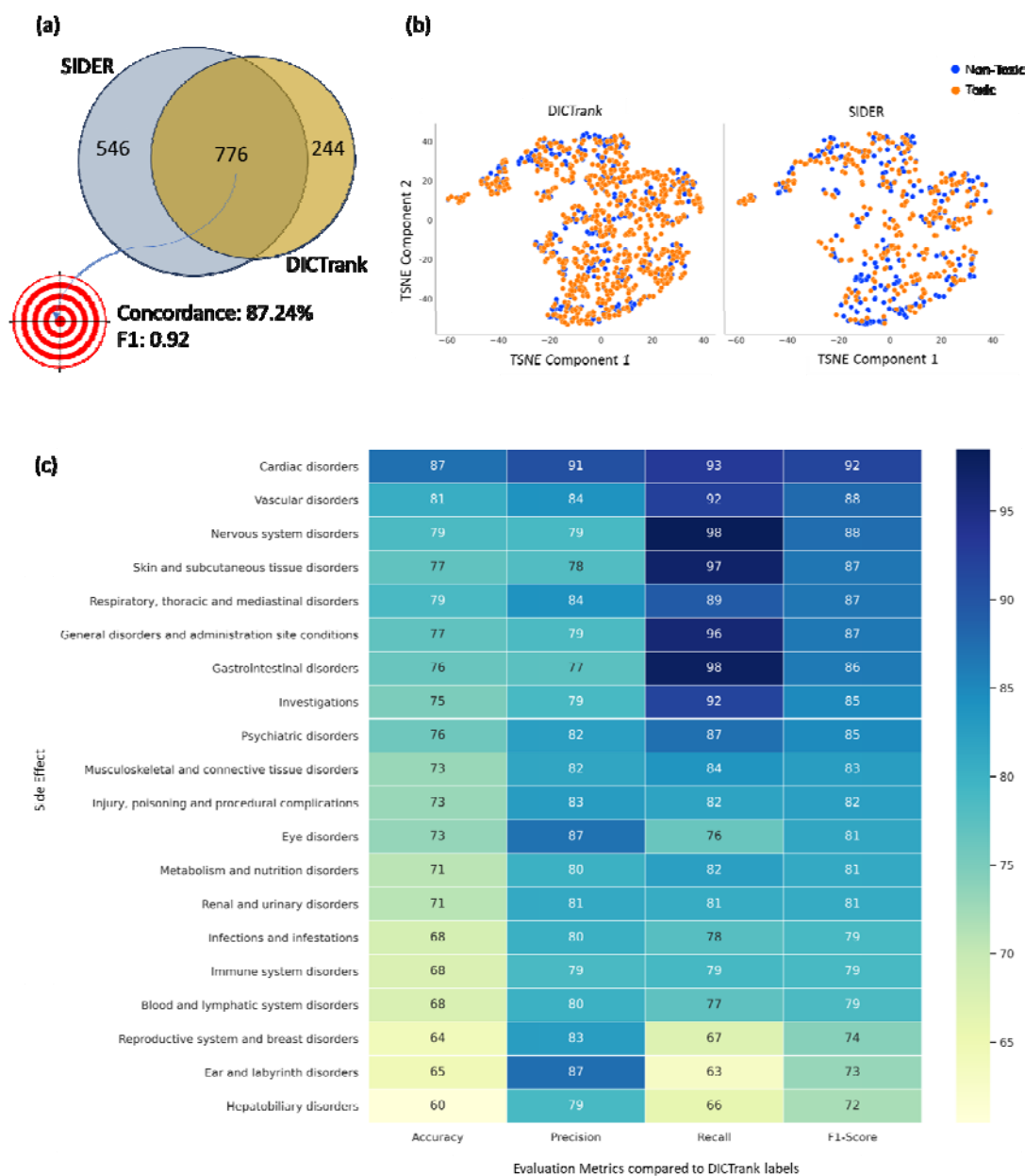


Figure 2: Comparison of the SIDER dataset with DICTrank: (a) the overlap and concordance (percentage of the total compounds with the same annotation) of DICTrank labels with SIDER cardiac disorder labels, (b) the overlay of SIDER and DICTrank chemical space in a principal component analysis using physicochemical properties, and (c) the positive predictive value of other side effects in SIDER for DICTrank labels (toxic/non-toxic).

Maximum total and unbound compound concentration in plasma predict cardiotoxicity

We next determined if a high C_{max} indicated compounds more likely to be cardiotoxic as seen in the case of doxorubicin where cardiotoxicity was found to be C_{max} dependent.⁵⁵ As

a single parameter, Cmax was not sufficiently discerning to differentiate between compounds that fall under the 'most-concern' and 'less-concern' categories as per the DICT concern classification (Figure 3). However, for both peak total plasma levels and peak unbound (active) plasma levels' Cmax, the median distributions were significantly distinguishable between cardiotoxic and non-toxic compounds (Figure 3) suggesting that Cmax can be a useful parameter in determining cardiotoxicity.

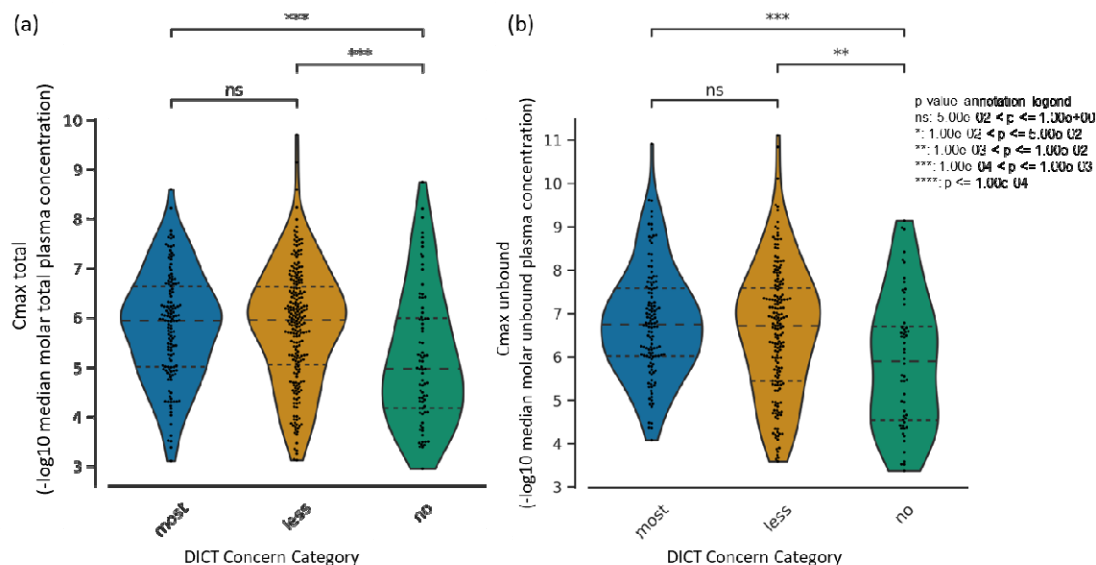


Figure 3: The distribution of (a) peak total concentration in plasma and (b) peak unbound (active) concentration in plasma for each drug in the DICTrank dataset across the three DICT concern categories.

Cyclooxygenase inhibition is predictive of cardiotoxicity concern

Turning to manual annotations of compound mechanisms of action and/or targets, we found that cyclooxygenase inhibitors⁵⁶ along with tyrosine kinase receptor inhibitors were the most significant annotations differentiating the various DICT concern categories (Table 3); this is plausible given cyclooxygenase inhibition, besides reducing inflammation, can also lead to increased blood pressure⁵⁷ while tyrosine kinase receptor inhibition can induce endoplasmic reticulum stress and inflammation in cardiomyocytes.⁵⁸ In agreement with this, known targets of prostaglandin endoperoxide synthases (PTGS1 and PTGS2 genes, which encode

cyclooxygenases COX-1 and COX-2) could significantly distinguish among most-, less- and no-DICT concern categories (Table 3).

CellScape-predicted protein targets such as hERG are predictive of cardiotoxicity

Among CellScape-predicted protein targets, the predicted activity of compounds against KCNH2 is best differentiated among the three DICT concern categories. The KCNH2 gene, also known as the human ether-à-go-go-related gene (hERG), is well known for its significance in the cardiac electrical cycle and hERG inhibition can lead to cardiac arrhythmias.⁵⁹ We also found that the top three features to distinguish the two DICTrank labels (cardiotoxic versus non-toxic) were α -l-fucosidase I, P-selectin, and carbonic anhydrase IX. The activity of plasma α -l-fucosidase has been pinpointed as a potential biomarker for cardiac hypertrophy and complements the currently used marker, atrial natriuretic peptide.⁶⁰ Elevated amounts of soluble P-selectin in the blood are evident in various heart-related conditions, like coronary artery disease, hypertension, and atrial fibrillation.⁶¹ Carbonic anhydrase IX plays a role in managing the intracellular pH in the heart muscle, a vital for the heart's functionality.⁶²

Table 3: Features from chemical and biological data sources that have a particularly high significance in the difference of distributions for the three DICT Concern Categories (Most, Less, No-concern).

Feature	P-value (statistical test)	Test applied	Feature space	Description/Biological interpretation	Source
Cmax (total)	2.97e-04	Mann Whitney Wilcoxon test two-sided with Bonferroni correction (most vs. no)	Pharmacokinetic parameters ³⁰	The peak total concentration of a drug in plasma indicates how much of the drug reaches the bloodstream.	55
Cmax (unbound)	7.58e-04			The peak unbound (active) concentration of a drug in plasma, indicates how much of the drug is available for interaction with its target.	
Cyclooxygenase inhibitor	2.98e-06	Chi-squared test	MOA (Drug Repurposing Hub ²⁹)	Inhibits cyclooxygenase enzymes, often leading to reduced inflammation but also increased blood pressure.	57
Tyrosine kinase receptor inhibitor	3.24e-05			Inhibition of tyrosine kinase receptors can affect cell growth and proliferation and can also induce endoplasmic reticulum stress, hypertension, heart failure, myocardial infarction, and cardiac arrhythmias.	58
PDGFR tyrosine kinase receptor inhibitor	3.24e-05				
PTGS2	9.67e-07	Chi-squared test	Known targets (Drug Repurposing Hub ²⁹)	Prostaglandin-endoperoxide synthase 2 (an enzyme) also known as COX-2.	57
PTGS1	9.67e-07			Prostaglandin-endoperoxide synthase 1 (an enzyme) also known as COX-1.	
HTR1D	1.27e-05			5-hydroxytryptamine receptor 1D, a serotonin receptor subtype. Previous enrichment analysis for methylation differences identified HTR1D among the genes with decreased promoter methylation, suggesting its involvement with serotonin receptors, which influences human cardiac	63,64

				function			
Q12809 at 100uM (KCNH2) (hERG)	1.21e-18	Kruskal-Wallis test Chi-squared test	CellScape Predicted Target ³⁴	The hERG gene encodes Kv11.1 channels crucial for heart function, linked to genetic and drug-induced arrhythmias	59		
Cytoplasm Granularity 2 ER	4.64e-04		Cell Painting ²⁴		Fine-grained smoothness of the ER staining. Disruptions in ER function can lead to ER stress, which is associated with various cardiovascular diseases.	65	
Cells Granularity 2 ER	1.05e-03					Contains information about the size, shape, number, or texture of nucleoli within the nucleus. This could encode signals for cellular stress.	66
Nuclei Texture Contrast RNA 3 0	1.19e-03						66
222103 at (ATF1)	1.01e-04		Gene Expression ^{19,25}		ATF1 is essential for cardiomyocyte function.	67	
201080 at (PIP4K2B)	6.29e-04					The role of PIP4k2 in cardiac disorders remains uncertain. PIP4Ks regulate insulin production and immune response, with PIP4k2c impacting TGFβ1 signaling which is vital in heart disease and other fibrotic conditions.	68
209092 s at (GLOD4)	8.14e-04						The physiological function of GLOD4 remains largely unexplored. The glyoxalase gene family, comprising six enzymes with roles in metabolism and disease prevention, is crucial for detoxifying reactive dicarbonyls and maintaining cellular homeostasis.
transport vesicle (GO:0030133)	5.53e-04		Gene Ontology ^{19,26}		Extracellular vesicles play important roles in cardiovascular communication, transporting bioactive molecules that both maintain heart health and contribute to cardiovascular diseases.	70	
negative regulation of potassium ion	1.04e-03					Cardiac K ⁺ channels play a crucial role in cardiac repolarization and their dysfunction can lead to arrhythmias.	71

transmembrane transport (GO:1901380)				
response to methylmercury (GO:0051597)	1.56e-03		Exposure to mercury (Hg) is considered to be an increased risk of developing cardiovascular system	⁷²
VSA EState6	6.67e-09	Physicochemical Descriptors from RDKit ³⁶	VSA EState Descriptor 6 (6.00 ≤ x < 6.07) related to molecular surface area and electronic state.	⁷³
Qed	1.20e-07		Quantitative estimate of drug-likeness, a measure indicating how drug-like a molecule is.	⁷⁴
NumHAcceptors	1.26e-07		Number of hydrogen bond acceptors in the molecule.	⁷⁵

Hypothesis-free omics data for cardiotoxicity are related to mechanisms of action

Omics data sources such as Cell Painting (imaging), gene expression, and Gene Ontology features cover a broad swath of biology, not specifically targeted to cardiac function. For Cell Painting, the fine-grained smoothness of the ER in the cytoplasm and RNA in the nucleus were the top features that differed significantly among toxicity classes. This is plausible given disruptions in ER function can lead to ER stress, which is associated with various cardiovascular diseases.⁶⁵ For the gene expression feature space, activating transcription factor 1 (ATF1), which is essential for cardiomyocyte function, was the top feature. The other two gene expression features that could distinguish DICT concern categories were phosphatidylinositol-5-phosphate 4-kinase type 2 beta (PIP4K2B) and glyoxalase domain containing 4 (GLOD4); both have indirect links to heart disease and other fibrotic conditions (Table 3). Among Gene Ontology annotations, we found that biological processes related to vesicle transport, potassium ion transmembrane transport, and response to methylmercury could best differentiate signals for concern categories. This is plausible given cardiomyocytes rely on vesicular transport for various functions, including the delivery of membrane proteins and lipids. The potassium ion channels play crucial roles in cardiac cell electrical activity and dysregulation can lead to arrhythmias and other heart complications.^{70,71} Exposure to mercury (Hg) is also considered a risk for ischemic heart disease.⁷²

Physicochemical Properties can differ among DICT concern categories

Among the various molecular descriptors evaluated in our study, VSA_EState6 could significantly distinguish among the DICT-concern categories. This electrotopological state descriptor aggregates the differences in electronegativity between an atom and its neighboring atoms in a molecule, adjusted by their relative distances while focusing on atoms with specific van der Waals surface area.⁷³ This suggests that specific electronic and spatial properties are captured by the VSA_EState6 descriptor, although difficult to interpret

directly. The second predictive feature, Qed, captures a quantitative estimation of the drug-likeness score that encapsulates the underlying distribution data for a range of drug properties.⁷⁴ The third predictive feature, NumHAcceptors refers to the number of hydrogen bond acceptors in the compound. Munawar et al. showed that the most potent hERG inhibitors typically possess two aromatic groups, one hydrophobic group, and one hydrogen bond acceptor, at specific relative distances from each other.⁷⁵

Structural alerts from DICTrank can detect compounds causing cardiac disorders from a retrospective analysis of DrugBank

We determined 59 structural alerts that distinguish cardiotoxic and non-toxic compounds in the DICTrank dataset (Figure 4). Two structural alerts had a high positive predictive value (PPV) for the DICT most-concern category, including one with aromatic rings. Aromatic rings can lead to π -stacking or hydrophobic interactions with aromatic rings of amino acids within the hERG channel cavity increasing the potential for blocking and subsequent cardiotoxic effects.⁷⁶ Six structural alerts distinguished toxic versus non-toxic compounds with a positive predictive value of 1 and more than ten occurrences in the dataset (the PPV was used to filter the structural alerts, hence is not an evaluation metric here). Structural alerts with tertiary amines were consistently protonated at physiological pH in the DICTrank dataset, suggesting their importance in biological activity and hERG channel binding.^{77,78} It is also known that compounds with secondary amine (more hydrogen bond donor number) are likely to be less potent hERG inhibitors compared to tertiary amine (less hydrogen bond donor number).⁷⁸

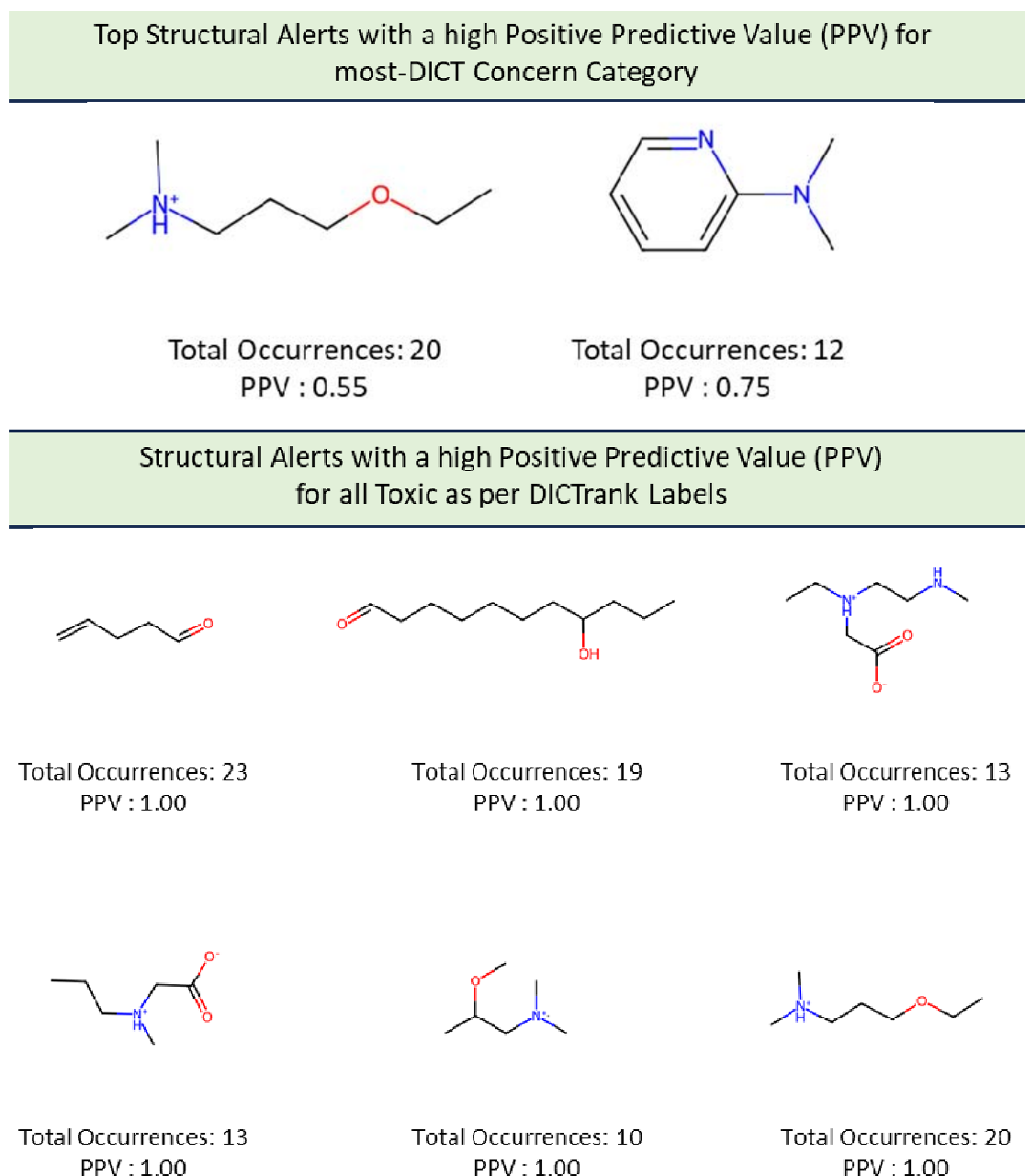
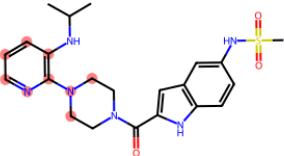
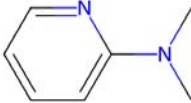
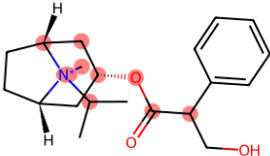
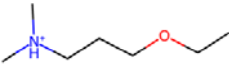
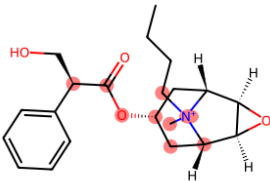
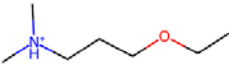
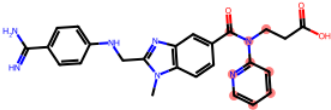
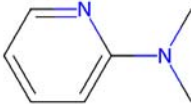


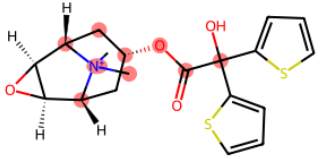
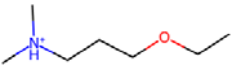
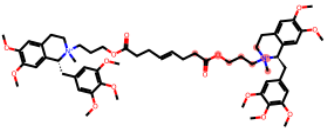
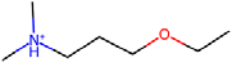
Figure 4: Structural alerts for (top) the most-concern DICT category and (bottom) DICTrank labels with more than ten occurrences and a PPV>0.6 for compounds in the DICTrank dataset.

We next analyzed compounds in DrugBank⁴⁷ for the presence of at least one of the two structural alerts above for the most-concern category. We annotated these hits with heart-related side effects from SIDER³² and their current status (approved, withdrawn, etc.) as indicated in DrugBank. We found six approved drugs, some experimental/investigational,

with reported cardiac disorders from SIDER (Table 4). These compounds spanned different classes of compounds, with the presence of a tertiary amine that remains protonated or aminopyridine rings as defined by the structural alerts. We found evidence in the literature for the risk of cardiovascular disorders for three of the six compounds, namely, ipratropium, tiotropium, and mivacurium.^{79–81} Overall, our analysis shows that the DICTrank dataset is a rich source of cardiotoxicity-causing compounds, with the potential to be used to build pharmacophore models and evaluate compounds with reported adverse events for their potential mechanisms of toxicity. Overall, we could detect multiple approved drugs that match the structural alerts for both the DICT most-concern category (as shown in Table 4) and for DICTrank labels for cardiotoxicity (further details in Supplementary Figure S2).

Table 4: Six hits from SIDER with the structural alerts for the DICT most-concern category.

Compounds Name	Compound Structure	Structural Alert Present	Status in DrugBank	Comment based on Literature
Delavirdine			Approved	
Ipratropium			Approved, Experimental	Use of ipratropium bromide within the past 6 months is associated with an increased risk of cardiovascular events. ⁷⁹
Butylscopolamine			Approved, Investigational	
Dabigatran			Approved, Investigational	Dabigatran effectively treats various cardiovascular diseases, with emerging studies assessing its use in malignancy associated VTE and other conditions. ⁸²

Tiotropium	 <p>The image shows the chemical structure of Tiotropium bromide. It features a tropane bicyclic core (8-azabicyclo[3.2.1]octane) with a bromine atom at the 8-position. Attached to the 7-position is a propyl chain, which is further substituted with a butyl chain and a 2-(thiophen-2-yl)ethyl group.</p>	 <p>The image shows the quaternary ammonium cation of Tiotropium, which is a nitrogen atom bonded to three methyl groups and one propyl chain that is further substituted with a butyl chain and a 2-(thiophen-2-yl)ethyl group.</p>	Approved	Tiotropium may increase the risk of coronary heart disease in patients over 55, especially high-exposure users. ⁸⁰
Mivacurium	 <p>The image shows the chemical structure of Mivacurium chloride. It consists of two tropane bicyclic cores (8-azabicyclo[3.2.1]octane) linked by a long, flexible polyether chain. Each tropane core has a methyl group at the 8-position and a propyl chain at the 7-position, which is further substituted with a butyl chain and a 2-(thiophen-2-yl)ethyl group.</p>	 <p>The image shows the quaternary ammonium cation of Mivacurium, which is a nitrogen atom bonded to three methyl groups and one propyl chain that is further substituted with a butyl chain and a 2-(thiophen-2-yl)ethyl group.</p>	Approved	Mivacurium chloride affects blood pressure and heart rate varying with dosage while rapid injections lead to pronounced changes. ⁸¹

Predictive models for DICTrank labels

Finally, given the promising signals seen in each data type as described above, we evaluated whether cardiotoxicity might be predicted using the data sources currently publicly available. Several data sources contained sufficient information to successfully train models to predict DICTrank labels (Table 2). We trained 11 models on four types of training data: the DICTrank compounds alone and DICTrank compounds enriched with cardiotoxic/non-toxic/all compounds in the SIDER dataset (as shown in Supplementary Table S1). A direct comparison of the predictive value of data sources is not possible due to the incomplete intersection of compounds with available data of each type. Still, we fixed the held-out test set of compounds to be those where data was available for all feature spaces such that only the training set of compounds varied among data sources. We trained two ensemble models, one using six models (structural, physicochemical, MOA, CellScape, MOA with Cmax total, and CellScape with Cmax total) that performed relatively well on the internal cross-validation (evaluation metrics from cross-validation for all feature space and dataset combinations, are given in Supplementary Table S4). This ensemble model was evaluated on an external test set of 90 compounds. Another ensemble model was built on all eleven models, which required testing on a smaller held-out test set due to the limited overlap of data. Evaluation metrics for all models are given in Supplementary Table S5.

Looking at each data source independently, we found that models using Mordred descriptors evaluated on the 90 compounds held-out test set (AUC: 0.84, AUCPR: 0.93; random AUC: 0.50, AUCPR: 0.72) performed better compared to models trained on predicted protein targets (AUC: 0.77, AUCPR: 0.89) and MOA annotations with Cmax (total) (AUC: 0.77, AUCPR: 0.90) (Figure 5a, b). In fact, models using Mordred descriptors were as good as the ensemble of six selected models (AUC: 0.83, AUCPR: 0.92; random AUCPR: 0.72) also evaluated on the 90 compounds held-out test set (Supplementary Figure S3). Further, models across most datasets performed with high AUCPR and F1 scores, with top-performing models using Mordred descriptors (AUCPR: 0.93; random AUCPR: 0.72) and

ensemble models (AUCPR: 0.93 for both ensemble models) when using the DICTrank dataset directly (Supplementary Figures S3a and S3b). Exceptions were models using the broad-based omics data - Cell Painting, Gene Expression, and Gene Ontology - where the performance was relatively poor and similar to random predictions according to the distribution of respective training data. This lack of predictive power may be inherent to the data sources but could also be due to the highly unbalanced and sparse training data available for these data sources (see Supplementary Table S2). When comparing the models evaluated with the smaller test set (Supplementary Figure S3), we found that models trained on the DICTrank dataset enriched with all SIDER compounds and using MOA data with Cmax (unbound) (AUCPR: 0.93, random AUCPR: 0.73) performed equally as the ensemble models that used predictions from all eleven models trained on just the DICTrank dataset (AUCPR: 0.93; random AUCPR: 0.73). Overall, a strong detection of cardiotoxicity was seen equally among the ensemble model and models using physicochemical descriptors.

We next analyzed the applicability domain of these models based on evaluating the quality of prediction for groups of compounds that are structurally dissimilar to the training data. We found that ensemble models and models using MOA annotations perform consistently well across the similarity range (Figure 5c). Models using Mordred descriptors, on the other hand, perform with slightly lower AUC-PR when compounds are structurally dissimilar to training data.

Finally, we predicted the DICTrank labels for 82 unique compounds that were labeled ambiguous in the original DICTrank dataset (Supplementary Table S6). We used Mordred descriptors and retrained the model on all 1020 compounds (training and held-out compounds) in DICTrank, except for the ambiguous compounds. We found that 43 of the 82 compounds were predicted to be cardiotoxic and 39 were predicted to be non-toxic and provided this list to the community for further study (Supplementary Data S6).

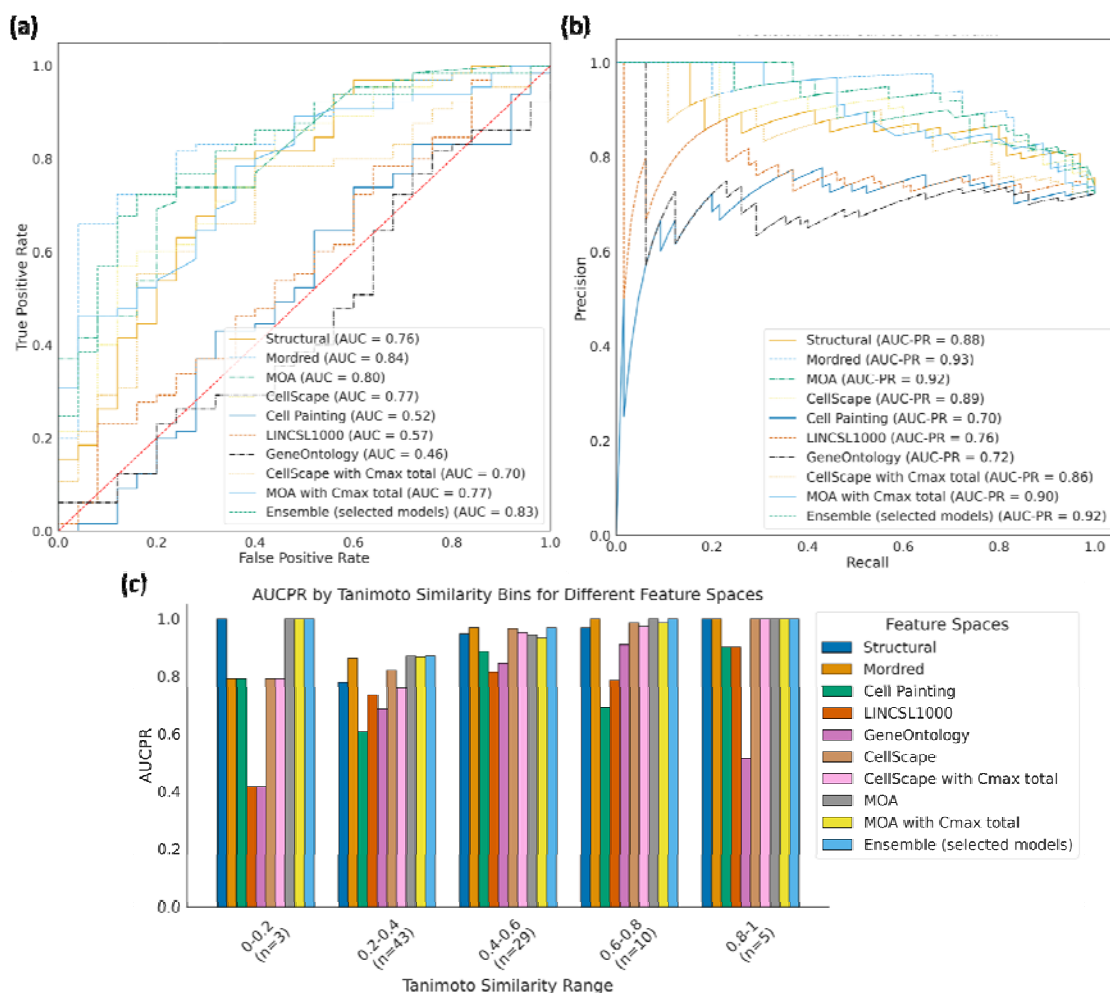


Figure 5: Comparison of evaluation metrics models built in this study with an external test set of 90 compounds evaluated by the (a) AUC-ROC and (b) AUCPR and (c) the performance of each model across compounds that are similar (dissimilar) to the training data. The ensemble mode in (c) is based on the models built on six data sources listed in the main text and evaluation is for 90 held-out compounds; results for the ensemble using 11 data sources and 78 held-out compounds are in Supplementary Figure S3.

Limitations of this Study

While we considered in this study various chemical and biological data sources, it is important to remember that conclusions are based on limited data. Certain feature spaces contain features that are computed based on chemical structure, such as CellScape target predictions and physicochemical properties, while datasets such as MOA and SIDER are manually gathered and have evidence of the presence and absence of evidence annotations. To train models using feature spaces such as Cell Painting, Gene Expression,

and Gene Ontology datasets, we dropped compounds where we could not find profiles (whether experimentally captured or imputed based on matching to highly similar compound profiles using the *v*-nn approach). The amount of training data (and also the class balance of SIDER/DICTrank labels) is lower for these models. Although we compare data sources using the same test compounds, the varying amounts of training data, and the differing types of compounds represented therein, can disadvantage some data sources versus others, such that we cannot with certainty compare the signal contained across the feature spaces. The poor performance of -omics data should therefore not yet be attributed to representing the signal in the feature space. Rather in this study, we aim to evaluate the signal present in the data that is available and build the best predictive models possible with public data. In the future, the availability of more data, for example, Cell Painting from JUMP-CP⁸³ and Recursion RxRx3⁸⁴ will significantly improve our ability to ascertain the presence of a signal for cardiotoxicity in -omics data.

Conclusions

In this work, we used biological data and chemical data (Figure 1) to predict drug-induced cardiotoxicity. We determined the feature contained in each data source that most differed between the most-concern versus non-toxic category for DICTrank and found these could drive mechanistic insights. Features from data sources such as predicted protein targets and annotated MOAs that could distinguish the DICT concern categories resembled activity against targets (ion channels in particular) that are mechanistically most plausible. We further evaluated these feature spaces using machine learning to build the first predictive models of DICTrank. Our findings indicate that models relying on physicochemical properties trained on larger training datasets performed on par with the ensemble models based on diverse data sources. The exploratory data analysis in this study suggests that as more -omics data becomes accessible in the future, it will enhance our ability to predict cardiotoxicity. Therefore, for the present, when constructing models using public datasets,

we advocate the use of Mordred descriptors and predicted targets (based on chemical structure), since these computed properties are readily available for compounds; they do not require experimental data and could be used to build models for cardiotoxicity. In the future, using biological data we can look into the biological pathways and mechanisms of DICT leading to better drug design and safer therapeutic strategies.

Associated Content

Supplemental Information. Supporting Information is available. Supporting Information (PDF). We released the Python code for our models which are publicly available at <https://github.com/srijitseal/DICTrank/> and all datasets at <https://doi.org/10.6084/m9.figshare.24312274.v1>

Acknowledgments

This work was performed using resources provided by the Cambridge Service for Data-Driven Discovery (CSD3) operated by the University of Cambridge Research Computing Service (www.csd3.cam.ac.uk), provided by Dell EMC and Intel using Tier-2 funding from the Engineering and Physical Sciences Research Council (capital grant EP/T022159/1), and DiRAC funding from the Science and Technology Facilities Council (www.dirac.ac.uk). Cartoons in TOC Figure and Figure 1 were created with DALLEv2 (<https://openai.com/dall-e-2>), Microsoft Designer (<https://designer.microsoft.com/>) and Bioicons (<https://bioicons.com>) which compiled images from the Database Centre for the life sciences/ TogoTV (<https://togotv.dbcls.jp>) and Servier (<https://smart.servier.com>).

Funding

S Seal acknowledges funding from the Cambridge Centre for Data-Driven Discovery (C2D3) and Accelerate Programme for Scientific Discovery. AEC, S Singh, and S Seal acknowledge funding from the National Institutes of Health (R35 GM122547 to AEC). OS acknowledges

funding from the Swedish Research Council (grants 2020-03731 and 2020-01865), FORMAS (grant 2022-00940), Swedish Cancer Foundation (22 2412 Pj 03 H), and Horizon Europe grant agreement #101057014 (PARC) and #101057442 (REMEDIA4ALL).

Author Declarations

S Singh and AEC serve as scientific advisors for companies that use image-based profiling and Cell Painting (AEC: Recursion, SyzOnc, S Singh: Waypoint Bio, Dewpoint Therapeutics) and receive honoraria for occasional talks at pharmaceutical and biotechnology companies. OS declares shares in Phenaros Pharmaceuticals. LGH is an employee at Ignota Labs where CellScape is a proprietary software. All other authors declare no relevant competing interests.

Author contributions

All authors have approved the final version of the manuscript. S Seal designed and performed exploratory data analysis, and implemented, and trained the models. LHG trained CellScape for the prediction of drug targets. S Seal, AEC, and S Singh analyzed the biological interpretation of morphological features. S Seal, OS, and AB analyzed the performance of machine learning models. S Seal wrote the manuscript with extensive discussions with all authors. All the authors (S Seal, OS, LHG, MGO, AEC, AB, and S Singh) reviewed, edited, and contributed to discussions on the manuscript.

References

1. Varga, Z.V., Ferdinandy, P., Liaudet, L., and Pacher, P. (2015). Drug-induced mitochondrial dysfunction and cardiotoxicity. *Am. J. Physiol. Heart Circ. Physiol.* 309, H1453–H1467. 10.1152/ajpheart.00554.2015.
2. Onakpoya, I.J., Heneghan, C.J., and Aronson, J.K. (2016). Post-marketing withdrawal of 462 medicinal products because of adverse drug reactions: a systematic review of the world literature. *BMC Med.* 14, 10. 10.1186/s12916-016-0553-2.
3. Dykens, J.A., and Will, Y. (2007). The significance of mitochondrial toxicity testing in drug development. *Drug Discov. Today* 12, 777–785. 10.1016/j.drudis.2007.07.013.
4. Clements, M., Millar, V., Williams, A.S., and Kalinka, S. (2015). Bridging Functional and Structural Cardiotoxicity Assays Using Human Embryonic Stem Cell-Derived Cardiomyocytes for a More Comprehensive Risk Assessment. *Toxicol. Sci.* 148, 241–260. 10.1093/toxsci/kfv180.
5. Mamoshina, P., Rodriguez, B., and Bueno-Orovio, A. (2021). Toward a broader view of mechanisms of drug cardiotoxicity. *Cell Rep Med* 2, 100216. 10.1016/j.xcrm.2021.100216.
6. McGowan, J.V., Chung, R., Maulik, A., Piotrowska, I., Walker, J.M., and Yellon, D.M. (2017). Anthracycline Chemotherapy and Cardiotoxicity. *Cardiovasc. Drugs Ther.* 31, 63–75. 10.1007/s10557-016-6711-0.
7. Sanguinetti, M.C., and Tristani-Firouzi, M. (2006). hERG potassium channels and cardiac arrhythmia. *Nature* 440, 463–469. 10.1038/nature04710.
8. Subbiah, I.M., Lenihan, D.J., and Tsimberidou, A.M. (2011). Cardiovascular toxicity profiles of vascular-disrupting agents. *Oncologist* 16, 1120–1130. 10.1634/theoncologist.2010-0432.
9. Gilbert, G., Demydenko, K., Dries, E., Puertas, R.D., Jin, X., Sipido, K., and Roderick, H.L. (2020). Calcium Signaling in Cardiomyocyte Function. *Cold Spring Harb. Perspect. Biol.* 12. 10.1101/cshperspect.a035428.
10. Thomas, T.P., and Grisanti, L.A. (2020). The Dynamic Interplay Between Cardiac Inflammation and Fibrosis. *Front. Physiol.* 11, 529075. 10.3389/fphys.2020.529075.
11. Heinzerling, L., Ott, P.A., Hodi, F.S., Husain, A.N., Tajmir-Riahi, A., Tawbi, H., Pauschinger, M., Gajewski, T.F., Lipson, E.J., and Luke, J.J. (2016). Cardiotoxicity associated with CTLA4 and PD1 blocking immunotherapy. *J Immunother Cancer* 4, 50. 10.1186/s40425-016-0152-y.
12. Hartupee, J., and Mann, D.L. (2017). Neurohormonal activation in heart failure with reduced ejection fraction. *Nat. Rev. Cardiol.* 14, 30–38. 10.1038/nrcardio.2016.163.
13. Li, M.-Y., Peng, L.-M., and Chen, X.-P. (2022). Pharmacogenomics in drug-induced cardiotoxicity: Current status and the future. *Front Cardiovasc Med* 9, 966261. 10.3389/fcvm.2022.966261.
14. Qu, Y., Li, T., Liu, Z., Li, D., and Tong, W. (2023). DICTrank: The largest reference list of 1318 human drugs ranked by risk of drug-induced cardiotoxicity using FDA labeling. *Drug Discov. Today* 28, 103770. 10.1016/j.drudis.2023.103770.

15. Chen, M., Suzuki, A., Thakkar, S., Yu, K., Hu, C., and Tong, W. (2016). DILLrank: the largest reference drug list ranked by the risk for developing drug-induced liver injury in humans. *Drug Discov. Today* 21, 648–653. 10.1016/j.drudis.2016.02.015.
16. Bender, A., and Cortés-Ciriano, I. (2021). Artificial intelligence in drug discovery: what is realistic, what are illusions? Part 1: Ways to make an impact, and why we are not there yet. *Drug Discov. Today* 26, 511–524. 10.1016/j.drudis.2020.12.009.
17. Bassan, A., Alves, V.M., Amberg, A., Anger, L.T., Beilke, L., Bender, A., Bernal, A., Cronin, M.T.D., Hsieh, J.-H., Johnson, C., et al. (2021). In silico approaches in organ toxicity hazard assessment: Current status and future needs for predicting heart, kidney and lung toxicities. *Comput Toxicol* 20. 10.1016/j.comtox.2021.100188.
18. Svensson, F., Zoufir, A., Mahmoud, S., Afzal, A.M., Smit, I., Giblin, K.A., Clements, P.J., Mettetal, J.T., Pointon, A., Harvey, J.S., et al. (2018). Information-Derived Mechanistic Hypotheses for Structural Cardiotoxicity. *Chem. Res. Toxicol.* 31, 1119–1127. 10.1021/acs.chemrestox.8b00159.
19. Wang, Z., Clark, N.R., and Ma'ayan, A. (2016). Drug-induced adverse events prediction with the LINCS L1000 data. *Bioinformatics* 32, 2338–2345. 10.1093/bioinformatics/btw168.
20. Galeano, D., Li, S., Gerstein, M., and Paccanaro, A. (2020). Predicting the frequencies of drug side effects. *Nat. Commun.* 11, 4575. 10.1038/s41467-020-18305-y.
21. Wu, Z., Ramsundar, B., Feinberg, E.N., Gomes, J., Geniesse, C., Pappu, A.S., Leswing, K., and Pande, V. (2018). MoleculeNet: a benchmark for molecular machine learning. *Chem. Sci.* 9, 513–530. 10.1039/c7sc02664a.
22. Duran-Frigola, M., and Aloy, P. (2013). Analysis of chemical and biological features yields mechanistic insights into drug side effects. *Chem. Biol.* 20, 594–603. 10.1016/j.chembiol.2013.03.017.
23. Jamal, S., Ali, W., Nagpal, P., Grover, S., and Grover, A. (2019). Computational models for the prediction of adverse cardiovascular drug reactions. *J. Transl. Med.* 17, 171. 10.1186/s12967-019-1918-z.
24. Bray, M.-A., Gustafsdottir, S.M., Rohban, M.H., Singh, S., Ljosa, V., Sokolnicki, K.L., Bittker, J.A., Bodycombe, N.E., Dancík, V., Hasaka, T.P., et al. (2017). A dataset of images and morphological profiles of 30 000 small-molecule treatments using the Cell Painting assay. *Gigascience* 6, 1–5. 10.1093/gigascience/giw014.
25. Subramanian, A., Narayan, R., Corsello, S.M., Peck, D.D., Natoli, T.E., Lu, X., Gould, J., Davis, J.F., Tubelli, A.A., Asiedu, J.K., et al. (2017). A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell* 171, 1437–1452.e17. 10.1016/j.cell.2017.10.049.
26. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25, 25–29. 10.1038/75556.
27. Seal, S., Yang, H., Vollmers, L., and Bender, A. (2021). Comparison of Cellular Morphological Descriptors and Molecular Fingerprints for the Prediction of Cytotoxicity- and Proliferation-Related Assays. *Chem. Res. Toxicol.* 34, 422–437. 10.1021/acs.chemrestox.0c00303.

28. Seal, S., Yang, H., Trapotsi, M.-A., Singh, S., Carreras-Puigvert, J., Spjuth, O., and Bender, A. (2023). Merging bioactivity predictions from cell morphology and chemical fingerprint models using similarity to training data. *J. Cheminform.* *15*, 56. [10.1186/s13321-023-00723-x](https://doi.org/10.1186/s13321-023-00723-x).
29. Corsello, S.M., Bittker, J.A., Liu, Z., Gould, J., McCarren, P., Hirschman, J.E., Johnston, S.E., Vrcic, A., Wong, B., Khan, M., et al. (2017). The Drug Repurposing Hub: a next-generation drug library and information resource. *Nat. Med.* *23*, 405–408. [10.1038/nm.4306](https://doi.org/10.1038/nm.4306).
30. Smit, I.A., Afzal, A.M., Allen, C.H.G., Svensson, F., Hanser, T., and Bender, A. (2021). Systematic Analysis of Protein Targets Associated with Adverse Events of Drugs from Clinical Trials and Postmarketing Reports. *Chem. Res. Toxicol.* *34*, 365–384. [10.1021/acs.chemrestox.0c00294](https://doi.org/10.1021/acs.chemrestox.0c00294).
31. Liu, A., Seal, S., Yang, H., and Bender, A. (2023). Using chemical and biological data to predict drug toxicity. *SLAS Discov* *28*, 53–64. [10.1016/j.slasd.2022.12.003](https://doi.org/10.1016/j.slasd.2022.12.003).
32. Kuhn, M., Letunic, I., Jensen, L.J., and Bork, P. (2016). The SIDER database of drugs and side effects. *Nucleic Acids Res.* *44*, D1075–D1079. [10.1093/nar/gkv1075](https://doi.org/10.1093/nar/gkv1075).
33. Broad Institute of Harvard Repurposing related drug annotations. <https://repuhub.broadinstitute.org/repurposing>.
34. Ignota Labs Ignota Labs. <https://ignotalabs.ai/>.
35. Prediction of drug side effects <http://www.maayanlab.net/SEP-L1000/>.
36. Landrum, G. (2013). Rdkit documentation. Release 1, 4.
37. Swain, M. (2019). MolVS: Molecule Validation and Standardization, 2018.
38. [Sul dimorphite dl: Protonate Your SMILES! Mirror of https://git.durrantlab.pitt.edu/jdurrant/dimorphite_dl/ \(GitHub\)](https://git.durrantlab.pitt.edu/jdurrant/dimorphite_dl/).
39. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Müller, A., Nothman, J., Louppe, G., et al. (2012). Scikit-learn: Machine Learning in Python. *arXiv [cs.LG]*, 2825–2830.
40. Sander, T., Freyss, J., von Korff, M., and Rufener, C. (2015). DataWarrior: an open-source program for chemistry aware data visualization and analysis. *J. Chem. Inf. Model.* *55*, 460–473. [10.1021/ci500588j](https://doi.org/10.1021/ci500588j).
41. Moriwaki, H., Tian, Y.-S., Kawashita, N., and Takagi, T. (2018). Mordred: a molecular descriptor calculator. *J. Cheminform.* *10*, 4. [10.1186/s13321-018-0258-y](https://doi.org/10.1186/s13321-018-0258-y).
42. Mervin, L.H., Bulusu, K.C., Kalash, L., Afzal, A.M., Svensson, F., Firth, M.A., Barrett, I., Engkvist, O., and Bender, A. (2018). Orthologue chemical space and its influence on target prediction. *Bioinformatics* *34*, 72–79. [10.1093/bioinformatics/btx525](https://doi.org/10.1093/bioinformatics/btx525).
43. [PIDGINv4: PIDGINv4 \(GitHub\) https://github.com/BenderGroup/PIDGINv4](https://github.com/BenderGroup/PIDGINv4).
44. Daina, A., Michielin, O., and Zoete, V. (2019). SwissTargetPrediction: updated data and new features for efficient prediction of protein targets of small molecules. *Nucleic Acids Res.* *47*, W357–W364. [10.1093/nar/gkz382](https://doi.org/10.1093/nar/gkz382).
45. Ferrari, T., Gini, G., Golbamaki Bakhtyari, N., and Benfenati, E. (2011). Mining toxicity

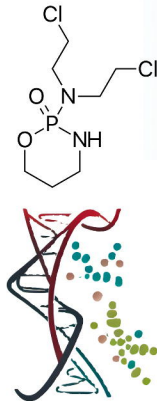
- structural alerts from SMILES: A new way to derive Structure Activity Relationships. In 2011 IEEE Symposium on Computational Intelligence and Data Mining (CIDM) (IEEE). 10.1109/cidm.2011.5949444.
46. Hemmerich, J., Troger, F., Füzi, B., and Ecker, G. (2020). Using Machine Learning Methods and Structural Alerts for Prediction of Mitochondrial Toxicity. *Mol. Inform.* *39*, e2000005. 10.1002/minf.202000005.
 47. Wishart, D.S., Feunang, Y.D., Guo, A.C., Lo, E.J., Marcu, A., Grant, J.R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., et al. (2018). DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* *46*, D1074–D1082. 10.1093/nar/gkx1037.
 48. Statistical functions (scipy.stats) — SciPy v1.11.3 Manual <https://docs.scipy.org/doc/scipy/reference/stats.html>.
 49. Seal, S., Carreras-Puigvert, J., Trapotsi, M.-A., Yang, H., Spjuth, O., and Bender, A. (2022). Integrating cell morphology with gene expression and chemical structure to aid mitochondrial toxicity detection. *Commun Biol* *5*, 858. 10.1038/s42003-022-03763-5.
 50. Lombardo, F., Berellini, G., and Obach, R.S. (2018). Trend Analysis of a Database of Intravenous Pharmacokinetic Parameters in Humans for 1352 Drug Compounds. *Drug Metab. Dispos.* *46*, 1466–1477. 10.1124/dmd.118.082966.
 51. Pang, L., Sager, P., Yang, X., Shi, H., Sannajust, F., Brock, M., Wu, J.C., Abi-Gerges, N., Lyn-Cook, B., Berridge, B.R., et al. (2019). Workshop Report: FDA Workshop on Improving Cardiotoxicity Assessment With Human-Relevant Platforms. *Circ. Res.* *125*, 855–867. 10.1161/CIRCRESAHA.119.315378.
 52. Rogers, D., and Hahn, M. (2010). Extended-Connectivity Fingerprints. *J. Chem. Inf. Model.* *50*, 742–754. 10.1021/ci100050t.
 53. Seal, S., Carreras-Puigvert, J., Carpenter, A.E., Spjuth, O., and Bender, A. (2023). From Pixels to Phenotypes: Integrating Image-Based Profiling with Cell Health Data Improves Interpretability. *bioRxiv*, 2023.07.14.549031. 10.1101/2023.07.14.549031.
 54. Slade, W.R., Jr, McNeal, A.C., and Tse, P.S. (1989). Neurologic complications of cardiovascular diseases. *J. Natl. Med. Assoc.* *81*, 193–197.
 55. Ishisaka, T., Kishi, S., Okura, K., Horikoshi, M., Yamashita, T., Mitsuke, Y., Shimizu, H., and Ueda, T. (2006). A precise pharmacodynamic study showing the advantage of a marked reduction in cardiotoxicity in continuous infusion of doxorubicin. *Leuk. Lymphoma* *47*, 1599–1607. 10.1080/10428190600580767.
 56. Patrono, C. (2016). Cardiovascular effects of cyclooxygenase-2 inhibitors: a mechanistic and clinical perspective. *Br. J. Clin. Pharmacol.* *82*, 957–964. 10.1111/bcp.13048.
 57. Singh, G., Miller, J.D., Huse, D.M., Pettitt, D., D’Agostino, R.B., and Russell, M.W. (2003). Consequences of increased systolic blood pressure in patients with osteoarthritis and rheumatoid arthritis. *J. Rheumatol.* *30*, 714–719.
 58. Wang, H., Wang, Y., Li, J., He, Z., Boswell, S.A., Chung, M., You, F., and Han, S. (2023). Three tyrosine kinase inhibitors cause cardiotoxicity by inducing endoplasmic reticulum stress and inflammation in cardiomyocytes. *BMC Med.* *21*, 147. 10.1186/s12916-023-02838-2.

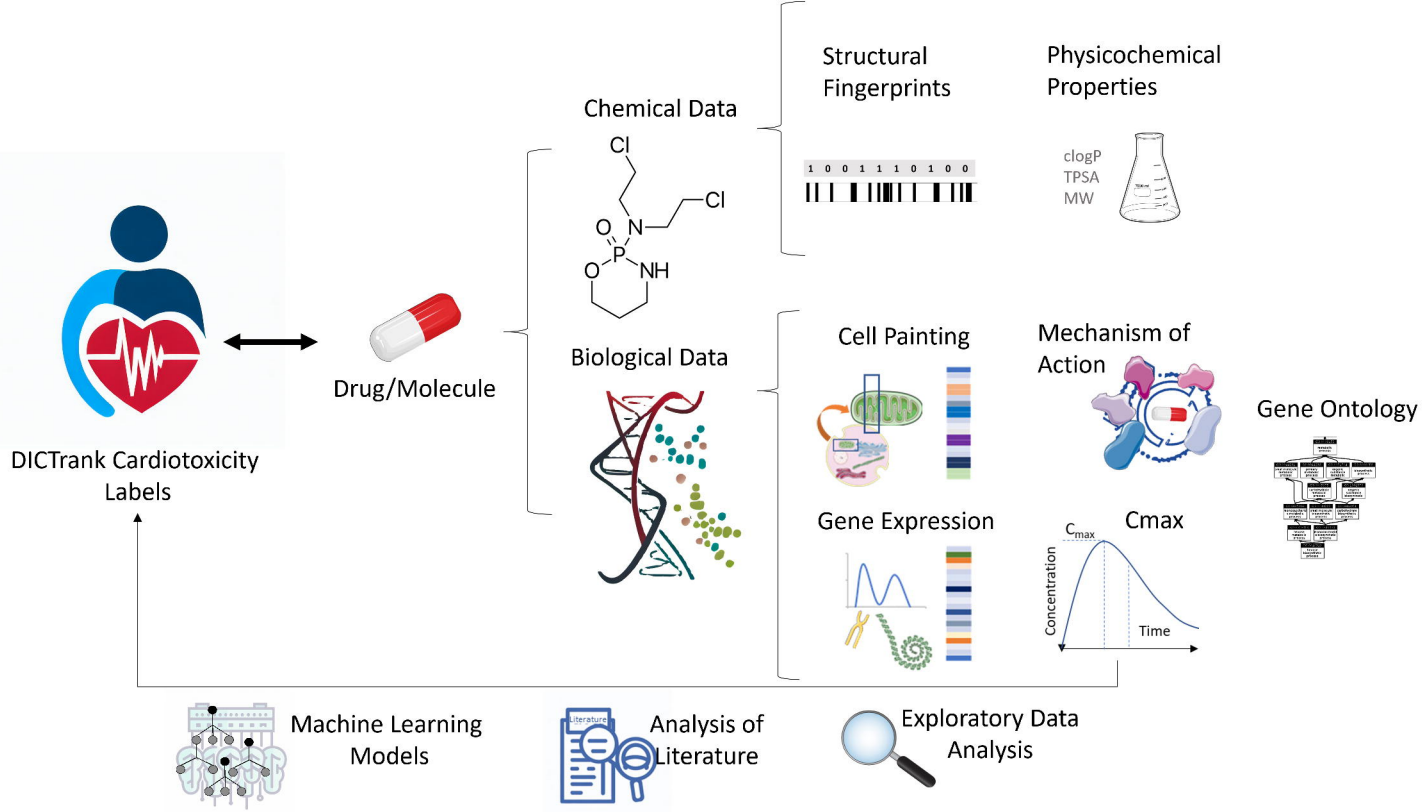
59. Vandenberg, J.I., Perry, M.D., Perrin, M.J., Mann, S.A., Ke, Y., and Hill, A.P. (2012). hERG K(+) channels: structure, function, and clinical significance. *Physiol. Rev.* *92*, 1393–1478. 10.1152/physrev.00036.2011.
60. Nagai-Okatani, C., and Minamino, N. (2016). Aberrant Glycosylation in the Left Ventricle and Plasma of Rats with Cardiac Hypertrophy and Heart Failure. *PLoS One* *11*, e0150210. 10.1371/journal.pone.0150210.
61. Blann, A.D., Nadar, S.K., and Lip, G.Y.H. (2003). The adhesion molecule P-selectin and cardiovascular disease. *Eur. Heart J.* *24*, 2166–2179. 10.1016/j.ehj.2003.08.021.
62. Nolly, M.B., Vargas, L.A., Correa, M.V., Lofeudo, J.M., Pinilla, A.O., Rueda, J.O.V., Guerrero-Gimenez, M.E., Swenson, E.R., Damiani, M.T., and Alvarez, B.V. (2021). Carbonic anhydrase IX and hypoxia-inducible factor 1 attenuate cardiac dysfunction after myocardial infarction. *Pflugers Arch.* *473*, 1273–1285. 10.1007/s00424-021-02592-5.
63. Bain, C.R., Ziemann, M., Kaspi, A., Khan, A.W., Taylor, R., Trahair, H., Khurana, I., Kaipananickal, H., Wallace, S., El-Osta, A., et al. (2020). DNA methylation patterns from peripheral blood separate coronary artery disease patients with and without heart failure. *ESC Heart Fail* *7*, 2468–2478. 10.1002/ehf2.12810.
64. Neumann, J., Hofmann, B., Dhein, S., and Gergs, U. (2023). Cardiac Roles of Serotonin (5-HT) and 5-HT-Receptors in Health and Disease. *Int. J. Mol. Sci.* *24*. 10.3390/ijms24054765.
65. Belmadani, S., and Matrougui, K. (2019). Broken heart: A matter of the endoplasmic reticulum stress bad management? *World J. Cardiol.* *11*, 159–170. 10.4330/wjc.v11.i6.159.
66. Hariharan, N., and Sussman, M.A. (2014). Stressing on the nucleolus in cardiovascular disease. *Biochim. Biophys. Acta* *1842*, 798–801. 10.1016/j.bbadis.2013.09.016.
67. Matus, M., Schulte, J., Sur, H., Seidl, M., Schütz, G., Schmitz, W., and Müller, F.U. (2008). Knockout of ATF1 leads to enhanced cardiac contractility and output. *FASEB J.* *22*. 10.1096/fasebj.22.1_supplement.1155.14.
68. Magadum, A., Singh, N., Kurian, A.A., Sharkar, M.T.K., Sultana, N., Chepurko, E., Kaur, K., Zak, M.M., Hadas, Y., Lebeche, D., et al. (2021). Therapeutic Delivery of Pip4k2c-Modified mRNA Attenuates Cardiac Hypertrophy and Fibrosis in the Failing Heart. *Adv. Sci.* *8*, 2004661. 10.1002/advs.202004661.
69. Farrera, D.O., and Galligan, J.J. (2022). The Human Glyoxalase Gene Family in Health and Disease. *Chem. Res. Toxicol.* *35*, 1766–1776. 10.1021/acs.chemrestox.2c00182.
70. Fu, S., Zhang, Y., Li, Y., Luo, L., Zhao, Y., and Yao, Y. (2020). Extracellular vesicles in cardiovascular diseases. *Cell Death Discov* *6*, 68. 10.1038/s41420-020-00305-y.
71. Grandi, E., Sanguinetti, M.C., Bartos, D.C., Bers, D.M., Chen-Izu, Y., Chiamvimonvat, N., Colecraft, H.M., Delisle, B.P., Heijman, J., Navedo, M.F., et al. (2017). Potassium channels in the heart: structure, function and regulation. *J. Physiol.* *595*, 2209–2228. 10.1113/JP272864.
72. Hu, X.F., Lowe, M., and Chan, H.M. (2021). Mercury exposure, cardiovascular disease, and mortality: A systematic review and dose-response meta-analysis. *Environ. Res.* *193*, 110538. 10.1016/j.envres.2020.110538.

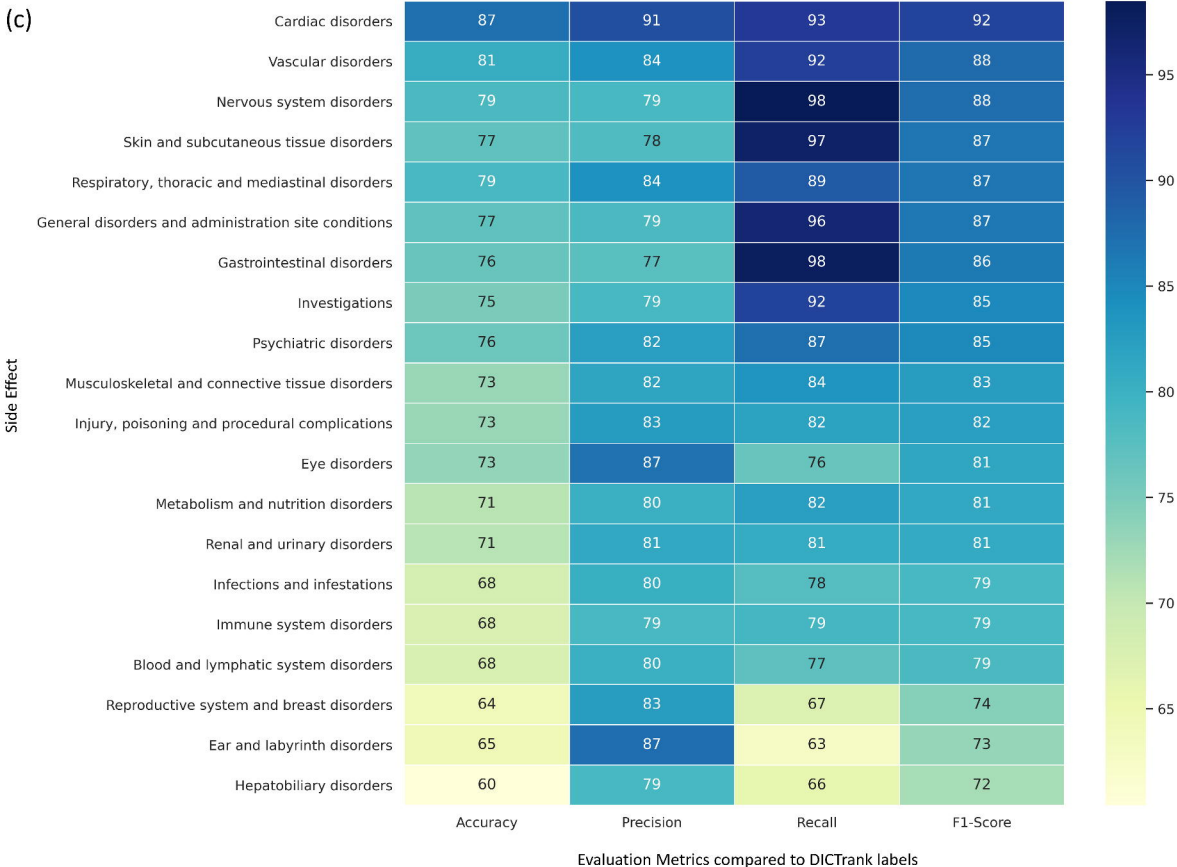
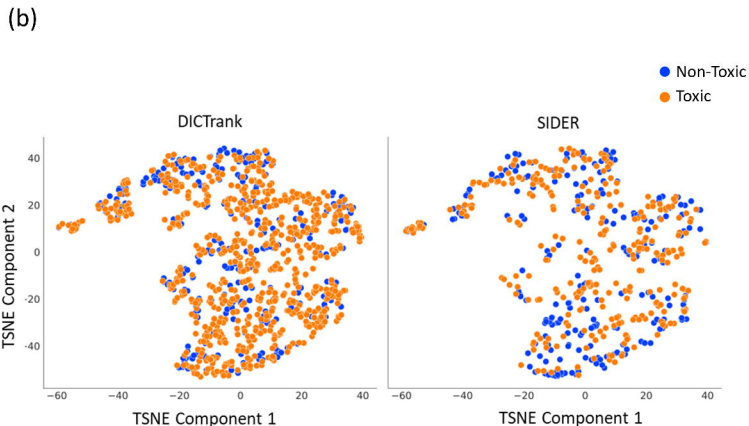
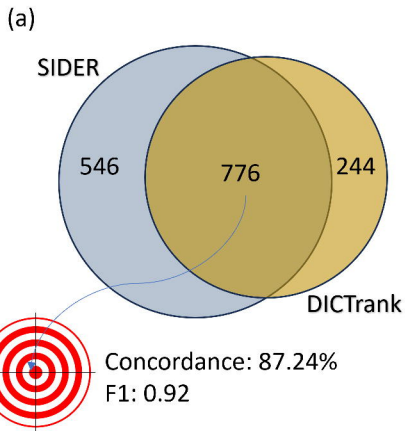
73. Kier, L.B., and Hall, L.H. (1990). An electrotopological-state index for atoms in molecules. *Pharm. Res.* 7, 801–807. 10.1023/a:1015952613760.
74. Bickerton, G.R., Paolini, G.V., Besnard, J., Muresan, S., and Hopkins, A.L. (2012). Quantifying the chemical beauty of drugs. *Nat. Chem.* 4, 90–98. 10.1038/nchem.1243.
75. Munawar, S., Windley, M.J., Tse, E.G., Todd, M.H., Hill, A.P., Vandenberg, J.I., and Jabeen, I. (2018). Experimentally Validated Pharmacoinformatics Approach to Predict hERG Inhibition Potential of New Chemical Entities. *Front. Pharmacol.* 9, 1035. 10.3389/fphar.2018.01035.
76. Polak, S., Wiśniowska, B., and 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, 183–206. 10.1002/jat.1395.
77. Cavalli, A., Poluzzi, E., De Ponti, F., and Recanatini, M. (2002). Toward a pharmacophore for drugs inducing the long QT syndrome: insights from a CoMFA study of HERG K(+) channel blockers. *J. Med. Chem.* 45, 3844–3853. 10.1021/jm0208875.
78. Garrido, A., Lepailleur, A., Mignani, S.M., Dallemagne, P., and Rochais, C. (2020). hERG toxicity assessment: Useful guidelines for drug design. *Eur. J. Med. Chem.* 195, 112290. 10.1016/j.ejmech.2020.112290.
79. Ogale, S.S., Lee, T.A., Au, D.H., Boudreau, D.M., and Sullivan, S.D. (2010). Cardiovascular events associated with ipratropium bromide in COPD. *Chest* 137, 13–19. 10.1378/chest.08-2367.
80. Shin, J., and Lee, J.H. (2022). Effects of tiotropium on the risk of coronary heart disease in patients with COPD: a nationwide cohort study. *Sci. Rep.* 12, 16674. 10.1038/s41598-022-21038-1.
81. Savarese, J.J., Ali, H.H., Basta, S.J., Scott, R.P., Embree, P.B., Wastila, W.B., Abou-Donia, M.M., and Gelb, C. (1989). The cardiovascular effects of mivacurium chloride (BW B1090U) in patients receiving nitrous oxide-opiate-barbiturate anesthesia. *Anesthesiology* 70, 386–394. 10.1097/0000542-198903000-00003.
82. Javed, A., Ajmal, M., and Wolfson, A. (2021). Dabigatran in cardiovascular disease management: A comprehensive review. *World J. Cardiol.* 13, 710–719. 10.4330/wjc.v13.i12.710.
83. Chandrasekaran, S.N., Ackerman, J., Alix, E., Michael Ando, D., Arevalo, J., Bennion, M., Boisseau, N., Borowa, A., Boyd, J.D., Brino, L., et al. (2023). JUMP Cell Painting dataset: morphological impact of 136,000 chemical and genetic perturbations. *bioRxiv*, 2023.03.23.534023. 10.1101/2023.03.23.534023.
84. Fay, M.M., Kraus, O., Victors, M., Arumugam, L., Vuggumudi, K., Urbanik, J., Hansen, K., Celik, S., Cernek, N., Jagannathan, G., et al. (2023). RxRx3: Phenomics Map of Biology. *bioRxiv*, 2023.02.07.527350. 10.1101/2023.02.07.527350.

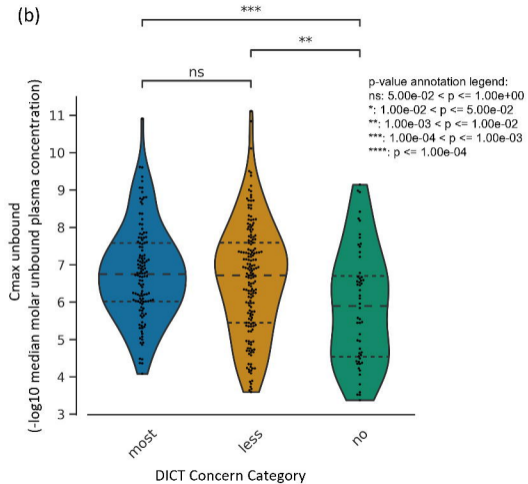
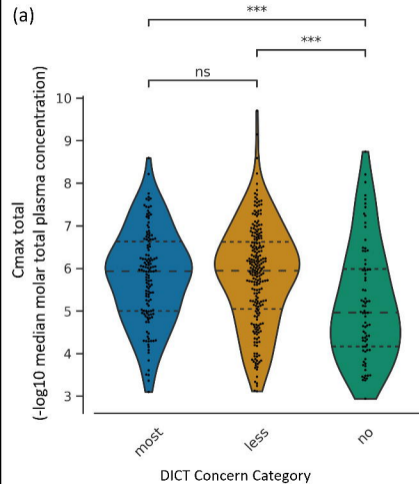


DICTrank Predictor

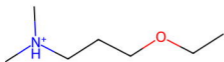




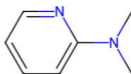




Top Structural Alerts with a high Positive Predictive Value (PPV) for most-DICT Concern Category



Total Occurrences: 20
PPV : 0.55

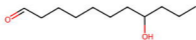


Total Occurrences: 12
PPV : 0.75

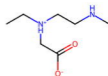
Structural Alerts with a high Positive Predictive Value (PPV) for all Toxic as per DICTrank Labels



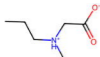
Total Occurrences: 23
PPV : 1.00



Total Occurrences: 19
PPV : 1.00



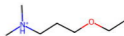
Total Occurrences: 13
PPV : 1.00



Total Occurrences: 13
PPV : 1.00



Total Occurrences: 10
PPV : 1.00



Total Occurrences: 20
PPV : 1.00

