Information-Derived Mechanistic Hypotheses for Structural Cardiotoxicity

Fredrik Svensson\textsuperscript{1*}, Azedine Zoufir\textsuperscript{1}, Samar Mahmoud\textsuperscript{1}, Avid M. Afzal\textsuperscript{1}, Ines Smit\textsuperscript{1}, Kathryn A. Giblin\textsuperscript{1}, Peter J. Clements\textsuperscript{2}, Jerome T. Mettetal\textsuperscript{3}, Amy Pointon\textsuperscript{4}, James S. Harvey\textsuperscript{2}, Nigel Greene\textsuperscript{5}, Richard V. Williams\textsuperscript{6}, Andreas Bender\textsuperscript{1,*}

\textsuperscript{1}Centre for Molecular Informatics, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK

\textsuperscript{2}GlaxoSmithKline R&D Ltd, Park Road, Ware, Hertfordshire SG12 0DP, UK

\textsuperscript{3}AstraZeneca, Drug Safety and Metabolism, 35 Gatehouse Dr, Waltham, MA 02451, USA

\textsuperscript{4}Safety and ADME Translational Sciences, AstraZeneca, Cambridge Science Park, Milton Road, Cambridge CB4 0WG, UK

\textsuperscript{5}Predictive Compound Safety and ADME, AstraZeneca, 35 Gatehouse Drive, Waltham, Massachusetts 02451, USA

\textsuperscript{6}Lhasa Limited, Granary Wharf House, 2 Canal Wharf, Leeds LS11 5PY, UK

\textsuperscript{*} Corresponding author e-mail: ab454@cam.ac.uk, f.svensson@ucl.ac.uk
Table of Contents Graphics

![Diagram of a table showing the connection between ToxCast and Compound Screening Results through FDA Adverse Event Reporting System (FAERS)]
Abstract

Adverse events resulting from drug therapy can be a cause of drug withdrawal, reduced and or restricted clinical use, as well as a major economic burden for society. To increase the safety of new drugs, there is a need to better understand the mechanisms causing the adverse events. One way to derive new mechanistic hypotheses is by linking data on drug adverse events with the drugs’ biological targets. In this study we have used datamining techniques and mutual information statistical approaches to find associations between reported adverse events collected from the FDA Adverse Event Reporting System and assay outcomes from ToxCast, with the aim to generate mechanistic hypotheses related to structural cardiotoxicity (morphological damage to cardiomyocytes and/or loss of viability). Our workflow identified 22 adverse event-assay outcome associations. From these associations 10 implicated targets could be substantiated with evidence from previous studies reported in the literature. For two of the identified targets, we also describe a more detailed mechanism, forming putative adverse outcome pathways (AOPs) associated with structural cardiotoxicity. Our study also highlights the difficulties deriving these type of associations from the very limited amount of data available.
Introduction

Despite major efforts to avoid off-target effects and toxicity in the development of new drugs, adverse events (AE) resulting from drug treatment remain a major clinical and economic burden.\(^1\) Toxicity is also a major cause of drug attrition, severely hampering the productivity of drug discovery and development.\(^2\) A detailed understanding of the mechanisms leading to AEs is important to allow for the efficient development of safer drugs in the future.

The recognition of the need for a more holistic understanding of toxicity is also reflected in the recently increased interest in the mechanisms of AEs, often formalised in Adverse Outcome Pathways (AOPs). AOPs are developed as a standardised way to structure evidence leading to an adverse outcome over different layers of biology.\(^3\) An AOP starts with a molecular initiating event (MIE)\(^4\) and then follows the pathway through a number of key events until it results in the adverse outcome. The pathway typically starts at the molecular level with a chemical entity interacting with for example a protein; the effect of this MIE is then seen at a cellular level, the tissue level, organ level, and finally at an individual level.\(^5\) An AOP can incorporate different amounts of information and evidence depending on the stage of development, where a putative AOP might only include plausible links between events, a fully developed quantitative AOP includes information not only on how the different events links to one another in the pathway but also on the concentrations and time of exposure required to trigger each event.\(^6\)

The compilation of an AOP requires a clear hypothesis on what biological mechanisms are involved in the observed outcome. One way to generate such hypotheses for further evaluation is by mining available data resources and linking recorded biological activities of chemicals to the observed outcomes,\(^5\) and a number of studies have reported on the use of computational techniques to derive information for AOP development. Edwards \textit{et al.} used databases such as ToxCast HTS
assays, the Comparative Toxicogenomics Database (CTD), and TG-GATEs to derive computationally predicted AOPs. These studies used frequent item set mining to find associations between entries in the different data sources. Perkins et al. used mutual information and network-based methods to analyze gene expression changes induced by compound exposure to derive putative AOPs for modulation of the hypothalamus-pituitary-gonadal endocrine axis in fathead minnows. Still, large quantities of information are available in various heterogeneous data sources and additional studies are required both to develop new methods to productively link these information sources and to assemble information on additional putative AOPs.

In order to make connections between AE and biological mechanisms there is a need to access large amounts of data on patient treatments and the biological targets perturbed by drug compounds. The FDA Adverse Event Reporting System (FAERS) is one of the largest collections of reported AEs and is frequently used as a data source for studies of AEs. Although a very valuable resource, the data can be hard to utilize due to the lack of standardization of the reports. Several studies have looked at ways to standardize the data to make it amenable to informatics studies. Another potential problem when making links between drugs and their AEs is that it is difficult sometimes to separate drug-induced AEs and deterioration in patient health related to the underlying disease, this is usually referred to as ‘indication bias’. In addition, biases stemming from co-medications and patient demographics might influence the results.

One source of bioactivity data is the United States Environmental Protection Agency toxicity forecaster program (ToxCast). Within ToxCast a large number of chemicals have been tested in a broad range of biological assays in order to create datasets for predictive toxicology. These bioassays involve screening diverse sets of target-based and phenotypic endpoints. Although a valuable resource for data mining, there are also some issues associated with the data in ToxCast
when used for constructing mechanistic hypotheses leading to toxicity (some of these limitations will be addressed in future developments of ToxCast\textsuperscript{18}). Ideally, information about a drug’s interaction with a particular target at biologically meaningful concentrations and conditions would be desired in order to identify an association between the target and an AE. However, the ToxCast assays generally measure the transcription of a particular gene within whole cells, rather than a direct interaction with a target. This adds complicating factors, such as the relevance of a particular cell type to the toxicity endpoint under evaluation. Furthermore, the doses tested in the ToxCast assays are not necessarily in the same range as the clinical concentrations and the time of exposure is clearly not the same as that which would be experienced by patients. The assays also lack the full complexity of an in vivo system and do consider the effect of metabolism and other ADME-T properties. Taken together, it is to be expected that many compounds will show activity in ToxCast assays without displaying any effects in vivo.

Cardiotoxicity can affect all components and functions of the cardiovascular system, either directly or indirectly and can be functional (acute alteration of the electromechanical function of the heart) and or structural (morphological damage to cardiomyocytes and/or loss of viability) in nature, as previously defined by Laverty et al. (2011).\textsuperscript{19} Structural and functional cardiotoxicity can sometimes be challenging to separate since structural damage to the heart also often results in a reduced function and chronic heart dysfunction can result in structural damage. A well-known example of functional cardiotoxicity is the QT prolongation and associated arrhythmia resulting from the inhibition of the hERG potassium channel,\textsuperscript{20} whereas structural cardiotoxicity is typically associated with certain classes of drugs such as anthracyclins\textsuperscript{21}, kinase inhibitors,\textsuperscript{22} and serotonergic drugs\textsuperscript{23}. A major problem is that the damage to the heart can be difficult to detect before it manifests clinically, by which point, the changes might be both life-threatening and
irreversible. The clinical situation is further complicated by the fact that it can take many years from exposure for the toxicity to manifest. Currently the mechanisms behind structural cardiotoxicity are poorly understood,\textsuperscript{24,25} and a better understanding of the mechanisms of drug-induced structural cardiotoxicity have been recognised as an area of great need.\textsuperscript{26}

In this study we investigate a data mining-based approach to generate mechanistic hypotheses for toxicity, exemplified by an analysis of structural cardiotoxicity. Data from FAERS and ToxCast was analysed to identify associations between reported assay activities and AEs, and the identified associations were evaluated by an extensive literature survey. The results indicate that the computationally identified associations can serve as valuable starting points for the construction of mechanistic hypotheses leading to structural cardiotoxicity. This serves as an example how heterogenous data sources can be combined and queried to derive mechanistic hypotheses, something that also can be extended to other data sources and outcomes.

**Methods**

An overview of the workflow used in this study is shown in Figure 1.

*FAERS data*

Drug-AE pairs were extracted from the standardized version of FAERS reported by Wang et al.\textsuperscript{14}

One concern when working with reported AEs is that the AE might originate in an underlying pathology rather than from the medication. One option is to filter the data to reduce this kind of bias. However, because the overlap of FAERS and ToxCast is very limited, removing too many drugs would reduce our chances of uncovering novel associations. For our analysis we therefore decided to implement and test three ways of accounting for indication based on various stringencies: 1. The less stringent approach was to not filter for indications at all; 2. the moderate
stringency filter consisted of removing all patient reports for which the indication started with ‘HEART’, ‘CARDI’ or ‘MYOCARD’; 3. the most stringent approach removed all drugs in FAERS where the drugs were used to treat a cardiovascular condition that matched any preferred term falling under the system organ class "Cardiac disorders". The second filter was deemed the most productive (see Results and Discussion) and the remainder of the methods section refers to the results obtained using this filter.

Apart from the above described process to reduce indication bias, no other confounding factor was addressed in this study.

For the resulting drug – AE-pairs, we applied an established method in signal detection, reporting odds-ratio (ROR) and Fisher’s exact test as described in van Puijenbroek. Drug – AE-pairs with an ROR ≥ 2 and Fisher’s test p-value ≤ 10^-4 were extracted, yielding a total of 6,715 drug – AE-pairs matching 2,711 unique drugs and 132 unique AEs.

**Cross-referencing with ToxCast**

The compounds used in this study were retrieved from ToxCast phase I,II and E1K screening (retrieved from actor.epa.gov/dashboard/ on 26th November 2015). To associate drugs with ToxCast target/readout information, the compounds extracted from FAERS in the previous step were cross-referenced with the data in ToxCast by matching standardized InChI strings. For the overlapping compounds, binary assay activity data (using the ToxCast “activity call”) was extracted from the ToxCast database, generating a compound – assay response matrix. The final dataset contained 88 compounds with information from 551 different assays and linked to 132 different AEs. Overall, the available activity data was very sparse with only 8% of the matrix populated.

**Calculation of Mutual Information Between ToxCast and FAERS data**
Before the mutual information was calculated we removed assays and adverse events with low variance (scikit-learn VarianceTreashold 0.09), leaving 207 assays and 13 AEs (when calculating the lowest mutual information pairs low variance filtering was omitted). To compute the pairwise mutual information between assays and adverse events, compounds that have not been tested in a particular assay were, for the purpose of the computational analysis, treated the same as compounds inactive in that assay. This was done to obtain vectors with the same lengths for all assay/AE pairs.

The associations between assays and adverse events were calculated using the normalised mutual information function in scikit-learn\(^3^0\) (range 0-1). Mutual information, which has long been used in information theory and has recently found applications also in bioinformatics,\(^3^1\) measures how much information about one variable can be derived from another (mutual dependence), in this case how much information about the AEs can be derived from the ToxCast assays.

The mutual information was computed for each assay – AE pair, and for each AE we identified the top three assays with the highest mutual information values. These were further required to have at least one observation (drug) that was associated with both the AE and displayed assay activity, leaving 22 assay – AE-pairs.
Figure 1. Schematic of the workflow used to identify AE-target associations. These associations were subsequently substantiated by cross referencing with previous studies available in the literature.

Background distributions

To generate background distributions for each assay and adverse event, we computed the probability of observing the assay and the event. The computed probability vector was used to generate background distributions using the `rv_discrete` class from SciPy v0.19.1. The size parameter was set to 10,000 to generate 10,000 samples, representing 10,000 random profiles of activity across ToxCast assays.

Open Targets query

To examine to what extent the significant targets from this study have previously been linked to cardiovascular diseases, we queried each target in the Open Targets Platform against “Cardiovascular Disease”. The Open Targets Platform is a large-scale integration of data sources linking drug targets and diseases and is designed to rank entities based on currently available
evidence for the associations, as well as prioritize targets for further investigation. The types of evidence in the platform include data from single nucleotide polymorphisms, animal models, text-mining, targets of approved drugs, and affected pathways. The platform provides an association score between 0 and 1 for a given target-disease association to summarize the strength of the evidence for the association.

_Literature searches_

For literature support, the main target (as indicated in the ToxCast annotations) name was used in combination with terms such as “cardiotoxicity”, “cardiovascular”, “cardiomyocyte”, as well as name of the adverse event, to search PubMed for studies implicating the target in structural cardiotoxicity. Studies considered relevant were those reporting for the investigated target any of the following: modulation of activity in connection with cardiovascular disease; considered to be a marker for heart disease; role in ischemia/reperfusion repair or injury; shown to have pharmacological effect in cardiomyocytes; expressed in cardiomyocytes; and known to play a role in homeostasis.

_Results and Discussion_

_FAERS report filter_

One major bias in this study could be the reports of cardiovascular events for patients with a known cardiovascular indication and current treatment. To prevent this we applied a filter to remove reports in which compounds were given for a cardiovascular indication. To investigate the effects on filtering the data we evaluated three different approaches; one without any filter, another moderate filter, and the third a more stringent filter completely removing drugs that can be given for any cardiovascular indication. These approaches generated 58, 22, and 11 associations.
respectively (see Supporting Information). We elected to use the moderate filter, leaving 22 associations. Although not exhaustive, this filter removes some of the obvious biases while still retaining a large portion of the available data.

**Target-AE associations**

For the final dataset we performed a mutual information analysis to identify assay activities associated with the reported AEs. The analysis produced 22 associations between a ToxCast assay and an AE, which are listed in Table 1. The assay activities that were associated with an AE included 16 distinct protein targets and two cell-based readouts. The assays that measure direct interaction of compounds with the target are indicated in Table 1. The other assays measure gene expression, transcription factor activity, or cellular endpoints.

**Table 1.** The 22 associations derived through the mutual information analysis. Rows including assay readouts that could be linked to structural cardiotoxicity in the literature survey are highlighted in grey. The mutual information and Pearson correlation is shown for each association, a higher value indicates a stronger association.

<table>
<thead>
<tr>
<th>Adverse Event</th>
<th>ToxCast Assay*</th>
<th>Assay Target/Readout</th>
<th>Direct Target Interaction</th>
<th>Mutual Information</th>
<th>Pearson Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cardiac failure congestive</td>
<td>BSK_BE3C_PA1_dow</td>
<td>PAI-1</td>
<td></td>
<td>0.120</td>
<td>0.369</td>
</tr>
<tr>
<td>Cardiac failure congestive</td>
<td>BSK_4H_Eotaxin3_dow</td>
<td>CCL26</td>
<td></td>
<td>0.116</td>
<td>0.366</td>
</tr>
<tr>
<td>Left atrial dilatation</td>
<td>ATG_PXRE_CIS_up</td>
<td>NR1I2</td>
<td></td>
<td>0.110</td>
<td>0.332</td>
</tr>
<tr>
<td>Left atrial dilatation</td>
<td>NVS_NR_cAR</td>
<td>AR</td>
<td>✓</td>
<td>0.105</td>
<td>0.337</td>
</tr>
<tr>
<td>Mitral valve incompetence</td>
<td>NVS_MP_rPBR</td>
<td>TSPO</td>
<td>✓</td>
<td>0.104</td>
<td>0.338</td>
</tr>
<tr>
<td>Mitral valve prolapse</td>
<td>NVS_NR_bER</td>
<td>ERα</td>
<td>✓</td>
<td>0.096</td>
<td>0.322</td>
</tr>
<tr>
<td>Cardiac failure congestive</td>
<td>Tox21_PPARg_BLA_anagonist_ratio</td>
<td>PPARγ</td>
<td>✓</td>
<td>0.096</td>
<td>0.314</td>
</tr>
<tr>
<td>Left atrial dilatation</td>
<td>ATG_VDRE_CIS_up</td>
<td>VDR</td>
<td></td>
<td>0.096</td>
<td>0.323</td>
</tr>
<tr>
<td>Mitral valve incompetence</td>
<td>NVS_NR_cAR</td>
<td>AR</td>
<td>✓</td>
<td>0.084</td>
<td>0.320</td>
</tr>
<tr>
<td>Mitral valve prolapse</td>
<td>NVS_NR_hER</td>
<td>ERα</td>
<td>✓</td>
<td>0.082</td>
<td>0.299</td>
</tr>
<tr>
<td>Condition</td>
<td>Assay Details</td>
<td>Gene/Protein</td>
<td>MI Value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------</td>
<td>------------------------------------</td>
<td>--------------</td>
<td>----------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mitral valve incompetence</td>
<td>BSK_BE3C_IP10_down</td>
<td>CXCL10</td>
<td>0.081</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left ventricular hypertrophy</td>
<td>BSK_CASM3C_MCSF_down</td>
<td>CSF1</td>
<td>0.081</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pericardial effusion</td>
<td>NVS_ADME_hCYP2C19</td>
<td>CYP2C19</td>
<td>0.074</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left ventricular dysfunction</td>
<td>BSK_3C_Eselectin_down</td>
<td>SELE</td>
<td>0.067</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left ventricular dysfunction</td>
<td>BSK_3C_TissueFactor_down</td>
<td>Tissue factor</td>
<td>0.067</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiomegaly</td>
<td>Tox21_PPARg_BLA_Agonist_ratio</td>
<td>PPARγ</td>
<td>0.066</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiomegaly</td>
<td>ATG_PPRE_CIS_up</td>
<td>PPARα</td>
<td>0.066</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ischaemic cardiomyopathy</td>
<td>APR_HepG2_StressKinase_72h_up</td>
<td>Stress kinases</td>
<td>0.066</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left ventricular dysfunction</td>
<td>BSK_hDFCGF_TIMP1_down</td>
<td>TIMP-1</td>
<td>0.066</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ischaemic cardiomyopathy</td>
<td>APR_HepG2_CellCycle_Arrest_72h_dn</td>
<td>Cell phenotype</td>
<td>0.060</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aortic valve incompetence</td>
<td>APR_HepG2_StressKinase_72h_up</td>
<td>Stress kinases</td>
<td>0.053</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aortic valve incompetence</td>
<td>BSK_3C_TissueFactor_down</td>
<td>Tissue factor</td>
<td>0.053</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* For description of the ToxCast assays please refer to https://actorws.epa.gov/actorws/toxcast/v01/assays

The obtained mutual information values were generally quite low, but this is not a surprising outcome. Since we have used AEs reported for drugs on the market, a very strong correlation between a certain mechanism and the AE cannot be expected, as those probably would have excluded the drug from being marketed in the first place.34

Compounds that were associated with the specific AEs were on average more frequently active in the associated assay than compounds that were not associated with AEs. Figure 2 shows the hit rate, which is the fraction of compounds tested that are active in the assay, of the compounds behind each of the selected 22 assay-AE associations.
Figure 2. The hit rates (fraction of compounds tested that were active in the assay) of compounds behind each of the selected 22 assay-AE associations in Table 1. On average, the AE-associated compounds are more frequently active in these selected ToxCast assays.

We tested whether the hit rate in each of the individual assay was significantly higher for AE-associated compounds than the rest of the compounds. For all but two associations (Left atrial dilatation - NVS_NR_cAR and Pericardial effusion -NVS_ADME_hCYP2C19) the set of compounds associated with the AE was more frequently active in the associated assay, although only for 8 associations were the activity ratios significantly different (Fisher’s exact test, Holm–Šidák correction for multiple testing, p < 0.05, see Supporting Information).

Background distributions
To evaluate the strength of the associations we compared the mutual information values from the 22 associations to the maximum mutual information for the same AEs and any ToxCast assay in the randomly generated background distribution with the same information content as the original data (see Supporting Information). Figure 3 shows the mutual information of the 22 top assay-AE associations in this study versus the maximum mutual information observed for the same adverse events in the randomly generated background distribution. This analysis indicated that all the associations in Table 1 were stronger than what would be expected from random data.

**Figure 3.** Histogram of the mutual information values observed for the top 22 assay-AE associations in this study versus the maximum mutual information observed for the same adverse events and any ToxCast assay in the randomly generated background distribution.

*Open Targets query*
We also extracted the bottom ranked associations for each of the AEs based on the MI analysis (see Supporting Information). To evaluate the strength of the associations compared to the top associations we used Open Targets association scores between the identified targets and “Cardiovascular disease”. The average association score for the targets from the top associations was 0.4 while the targets from the bottom associations had an average score of 0.26 (Supporting Information). Figure 4 shows the Open Targets association scores for the top and bottom associated targets of each AE. More of the top-ranked targets have perfect scores of 1, indicating they have been strongly associated with cardiovascular diseases based on the integrated evidence in Open Targets. These targets are AR, PPARα, PPARγ, and VDR. For the other targets, the top and bottom scores are in the same range. While this could mean that the integrated evidence does not support the associations, it could also be due to less data being available for these targets and the associations being novel.
**Figure 4.** The Open Targets association scores for ‘Cardiovascular Disease’ of targets ranked highest and lowest by the mutual information analysis. The association score represents the strength of evidence for associations between targets and diseases based on currently available integrated data. The horizontal lines show the mean.

In summary, our computational validations appear to indicate that the generated associations could be informative of structural cardiotoxicity.

*Literature searches*

In order to further validate the findings, we conducted a survey of the available literature for evidence supporting the above associations (see methods section for details). We elected to look for associations between the targets and any form of structural cardiotoxicity, not only the specific AE indicated by the MI analysis. For 10 out of the 16 indicated targets we could find evidence supporting a link between the mechanism queried in the assay and structural cardiotoxicity (highlighted in Table 1). A summary of the reported mechanisms is shown in Table 2.

**Table 2.** Targets with literature evidence supporting an association to structural cardiotoxicity.

<table>
<thead>
<tr>
<th>Target</th>
<th>Biological Role Linked to Structural Cardiotoxicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERα</td>
<td>Cardioprotective role in mouse ischemia model[^{35}]</td>
</tr>
<tr>
<td>PAI-1</td>
<td>Related to myocardial neovascularization and cardiomyocyte apoptosis.[^{36}]</td>
</tr>
<tr>
<td>PPARα</td>
<td>Implicated in cardiomyocyte hypertrophy.[^{37,38}] Increased levels of PPARα may cause adverse remodeling of heart structure and metabolism.[^{39}] PPAR family is important to regulate cardiac metabolism.[^{40}]</td>
</tr>
<tr>
<td>PPARγ</td>
<td>Both PPARγ knock out mice and mice treated with PPARγ agonists develop cardiac hypertrophy.[^{41}] PPAR family is important to regulate cardiac metabolism.[^{40}]</td>
</tr>
<tr>
<td>AR</td>
<td>Protective effect in cardiomyocytes[^{42}]</td>
</tr>
<tr>
<td>Tissue factor</td>
<td>Linked to fibrosis[^{33}]</td>
</tr>
<tr>
<td>TIMP-1</td>
<td>Improves cardiac function on ischemic cardiomyopathy model rats.[^{44}]</td>
</tr>
<tr>
<td>TSPO</td>
<td>TSPO ligands inhibits toxicity in isolated cardiomyocytes.[^{45}] Related to reactive oxygen species (ROS).[^{46-49}]</td>
</tr>
<tr>
<td>VDR</td>
<td>Vitamin D deficiency is linked to cardiovascular disease, including left atrial /ventricle dysfunction.[^{50-52}]</td>
</tr>
</tbody>
</table>
CSF1 Improves cardiac function after ischemic injury by inducing vascular endothelial growth factor and cardiomyocyte survival.\textsuperscript{53,54} Indicated to be important for the progression of cardiomyopathy.\textsuperscript{55}

The assay readouts from ToxCast would typically detect either an increase or a decrease in the readout, however for our literature analysis we did not consider the directionality of the effect (i.e. stimulation or inhibition of the indicated target), but rather if any modulation of the target could be connected to structural cardiotoxicity. We chose this approach since the long-term effect of agonising or antagonising a certain receptor is often not known, and we reason that any type of effect disturbing a mechanism documented to be important for cardiotoxicity or cardio-protection might potentially result in toxicity.

Some of the identified targets were more evidently linked to structural cardiotoxicity than others in scientific literature. The most well-documented targets were tissue factor (TF) and translocation protein (18 kDa) TSPO, and the evidence for these mechanisms is described in more detail below.

*Tissue factor and structural cardiac disorders*

The normal, physiological role of TF is essential in the blood coagulation cascade.\textsuperscript{43} Upon injury to the blood vessel, TF detects activated Factor VII (FVII) and initiates the blood coagulation cascade via thrombin production, ultimately leading to a fibrinogen blood clot.\textsuperscript{43} In addition, non-coagulation functions of TF include the activation of protease-activated receptors (PAR) such as PAR-1, PAR-2 and PAR-3. These and related signaling pathways activated by TF result in pro-inflammatory, pro-angiogenic and anti-apoptotic effects,\textsuperscript{43} which are also vital in embryonic development.\textsuperscript{56} Cardiac muscle expresses high levels of TF compared to skeletal muscle, and within the heart TF levels are highest in the left ventricle.\textsuperscript{43}

There is evidence to support a role for dysregulated TF in clinical heart pathologies. For example, elevated circulating plasma TF has been linked to atherosclerotic cardiovascular disease
and acute coronary syndrome.\textsuperscript{43} Furthermore, TF expression was downregulated in myocardial biopsies from patients with hypertension and hypertension-induced ventricular hypertrophy.\textsuperscript{57} Similar downregulation of TF in the heart was found in patients with dilated cardiomyopathy, where TF levels correlated with the ejection fraction.\textsuperscript{58}

Some animal studies also provide evidence for the involvement of TF in pathological cardiac tissue remodeling, which is closely linked to the structural cardiotoxicity endpoint considered in the current work. A genetic variant of mice expressing very low levels of TF developed fibrosis selectively in the heart, and displayed impaired heart contractility and left ventricular dysfunction.\textsuperscript{59} Histological examination revealed that hemosiderin deposits, which are evidence of haemorrhage, were associated with areas of fibrosis.\textsuperscript{59} While low-TF mice have normal haemostasis under normal conditions, they appear to have an abnormal haemostatic response to vessel injuries, probably in part due to low thrombin.\textsuperscript{59} The authors propose that this impaired haemostasis results in haemorrhages from cardiac vessels and subsequent fibrosis in the heart.\textsuperscript{59}

Another suggested mechanism linking TF and cardiomyopathy involves the dysregulation of angiogenic factors influenced by TF.\textsuperscript{43} TF activation induces the expression of angiogenic factors such as vascular endothelial growth factor (VEGF).\textsuperscript{43,60} Thrombin, which is downstream of TF activation, also increases the expression of VEGF and angiopoietin-2 via PAR-1.\textsuperscript{43} These angiogenic factors are essential for vascular maintenance and repair processes, for instance in response to ischaemic injury in dogs.\textsuperscript{61} Several studies have linked imbalances of angiogenic factors to cardiomyopathy. For example, in a knock-out study of mice lacking PGC-1α, a coactivator of VEGF and other pro-angiogenic factors, the mice developed peripartum cardiomyopathy with left ventricular dysfunction, having ventricular dilation and reduced contractile function.\textsuperscript{57} In another study, mice lacking one isoform of VEGF developed ischaemic
cardiomyopathy due to left ventricular failure and dilated cardiomegaly.\textsuperscript{61} A potential pathway for low TF-induced heart failure derived from the literature evidence is shown in Figure 5, showing two suggested mechanisms.

\textbf{Figure 5.} Putative AOP connecting TF to heart failure \textit{via} left ventricular dysfunction. Because the ToxCast assay measured gene expression of TF instead of direct molecular interaction, it is possible that the actual molecular initiating event (MIE) is another interaction upstream of the downregulation of TF.

\textit{Translocation protein and structural cardiac disorders}

Another association identified was the one between the “NVS_MP_hPBR” assay and “mitral valve incompetence” (MVI). The “NVS_MP_hPBR” assay measures the changes to scintillation counting signals produced from the receptor-ligand binding of the key ligand $[^{3}\text{H}]-\text{PK11195}$, and are indicative of a change in transporter function and kinetics for the TSPO, whereas MVI can be defined as the dysfunction of the mitral valve closing, causing blood to leak back into the atrium after contraction.\textsuperscript{62}

TSPO, also called peripheral benzodiazepine receptor (PBR), is located in the outer membrane of mitochondria in various tissues including the heart and plays a role in cholesterol transportation into mitochondria.\textsuperscript{63} It has been shown that TSPO ligands affect the mitochondrial permeability pore (mPTP),\textsuperscript{64} and that inhibitors of TSPO delay its opening.\textsuperscript{65} Hence, mPTP activity seems to be directly linked to TSPO binding. Additionally, it has been shown that TSPO plays a role in the
production of reactive oxygen species (ROS).46–49 As a consequence of this oxidative stress, the opening of the mPTP66,67 is activated, which in turn leads to cardiomyocyte apoptosis.68,69 This in turn can lead to MVI since mitral valve incompetence can occur in patients with dilated cardiomyopathy in the absence of primary valvular disease: valve leakage is associated with annular dilatation.70 A potential pathway connecting TSPO with MVI is shown in Figure 6.

**Figure 6.** Putative AOP linking TSPO to MVI.

**Limitations**

The above case studies illustrate the complexity of the mechanisms potentially leading to structural cardiotoxicity. Further work is needed to understand the full pathway and to develop formalised AOPs for the identified associations.

Although we were able to find literature support for many of the identified targets (10/16), it is important to highlight some of the difficulties associated with the reported approach. The small number of compounds causing structural cardiotoxicity makes it difficult to establish strong signals that are statistically significant. This means that there is a great risk of pursuing false signals. Alternative approaches to validate the findings, such as through the literature, always risk suffering from confirmation bias. Only by accessing and including more data can this issue be eliminated.

Nevertheless, data mining of post-marketing safety data in combination with compound profiling information appears to be a promising way to identify associations between observed AEs and plausible mechanisms, as demonstrated here for structural cardiotoxicity as a case
study. The mechanistic hypotheses generated by this approach, while in themselves not definitive, can serve as promising starting points for further studies. For example, pharmaceutical companies use panels of known safety targets during preclinical drug development. Of the targets suggested in this study, AR, ER, and NR1I2 are on these published lists, but not specifically for cardiac associations. While more evidence for the associations suggested in this study is needed, the results from this study could be useful in selecting candidates for future inclusion in such safety studies.

**Conclusions**

Using the wealth of data available on AEs and compound biological effects can be a valuable complement to more traditional approaches for identifying mechanisms hypotheses leading to AEs. In this study several putative mechanisms associated with structural cardiotoxicity were mined from the FAERS and ToxCast databases. The analysis yielded 22 associations connecting a mechanism to structural cardiotoxicity. A detailed literature survey produced support for 10 of the identified associations, and for two of the associations, a large number of reports were available in the literature, allowing for the construction of putative mechanistic pathways connecting these targets with structural cardiotoxicity.

**Supporting Information**

Information on the compounds driving the associations, statistics and OpenTarget scores.

**Acknowledgements**
The authors thank Jonathan Vessey for helpful comments on the manuscript.

AMA gratefully acknowledges AstraZeneca for funding.

Abbreviations

Androgen receptor, AR; C-X-C motif chemokine 10, CXCL10; Chemokine (C-C motif) ligand 26, CCL26; colony stimulating factor 1, CSF1; Cytochrome P450 2C19, CYP2C19; Estrogen receptor α, ERα; nuclear receptor subfamily 1, group I, member 2, NR1I2; Peroxisome proliferator-activated receptor alpha, PPARα; Peroxisome proliferator-activated receptor gamma, PPARγ; Plasminogen activator inhibitor-1, PAI-1; selectin E, SELE; TIMP metallopeptidase inhibitor 1, TIMP1; Translocator protein, TSPO; Vitamin D receptor, VDR

References


Events in the US FDA’s Spontaneous Reports Database. *Drug Saf.* 25 (6), 381–392.


(57) Luther, T.; Dittert, D. D.; Kotzsch, M.; Erlich, J.; Albrecht, S.; Mackman, N.; Müller, M.


