Machine learning guided association of adverse drug reactions with *in vitro* target-based pharmacology

Seda Arat [1 # *], Robert Ietswaart [2 # *], Amanda X. Chen [3,4 #], Saman Farahmand [5 #], Bumjun Kim [6], Duncan Armstrong [7], Laszlo Urban [7 *]

# equal contributions

* correspondence:

seda8arat@gmail.com

robert_ietswaart@hms.harvard.edu

laszlo.urban@novartis.com

1. The Jackson Laboratory, Farmington, CT 06032
2. Department of Genetics, Blavatnik Institute, Harvard Medical School, Boston, MA 02115
3. David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02142
4. Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139
5. Department of Biology, University of Massachusetts Boston, Boston, MA 02125
6. Department of Chemical Engineering, Northeastern University, Boston, MA 02115
7. Novartis Institutes for Biomedical Research, Cambridge, MA 02139

Abstract

Adverse drug reactions (ADRs) are one of the leading causes of morbidity and mortality in health care. Understanding which drug targets are linked to ADRs can lead to the development of safer medicines. Here, we analyze *in vitro* secondary pharmacology of common (off) targets for 2134 marketed drugs. To associate these drugs with human ADRs, we utilized FDA Adverse Event Reports and developed random forest models that predict ADR occurrences from *in vitro* pharmacological profiles. By evaluating Gini importance scores of model features, we identify 221 target-ADR associations. Among these are established relations, such as the association of *in vitro* hERG binding with cardiac arrhythmias, which validate our machine learning approach. Evidence on bile acid metabolism supports our identification of associations between the Bile Salt Export Pump and renal, thyroid, lipid metabolism, respiratory tract and central nervous system disorders. Unexpectedly, our model suggests PDE3 is associated with 40 ADRs, including congenital renal disorders. These associations provide a comprehensive resource to support drug development and human biology studies.

Keywords

Adverse drug reactions, adverse event report, FAERS, secondary pharmacology, machine learning, statistical modeling, drug discovery & development, drug safety.

Toxicity is one of the major causes of termination, withdrawal or labeling of a drug candidate or drug, other than lack of efficacy [1–3]. There is an urgent need to better identify toxic on- and off-target effects on vital organ systems especially for cardiovascular, renal, hepatic and central nervous system (CNS)-related toxicities; furthermore, there is a desire to reduce cost and labor in preclinical assays and drug testing on non-human species [4–6]. *In vitro* pharmacological assays have been widely used to screen for possible off-targets and potential adverse effects and eliminate compounds that are not safe enough in the drug development stage as early as possible [5,7]. However, systematic prediction of compound safety and potential adverse events associated with a compound is still a challenge for the pharmaceutical industry.

Artificial intelligence can be very insightful for many different stages of drug discovery and development, such as automation in pharmacology assays, clinical trials, and basic science research. Previous studies have focused on predicting structure-function relationships based on chemical structure of small molecules and potency assays that probe the physicochemical properties of compounds to estimate associations with off-targets [8]. However, the diversity of structures that interact with targets, even when they are well described like human Ether-a-go-go-related gene (hERG), make it challenging to produce reliable models [9]. Several papers provide small, hand-curated databases providing up to 70 pharmacological targets (i.e. receptors, ion channels, transporters, etc.) with established links to adverse side effects based on a scientific literature search [5,7,10,11]. Mirams et al. recently described how integration of data from multiple ion channels (e.g. hERG, sodium, L-type calcium) provided improved *in silico* prediction of torsadogenic risk [12]. Chen et al. proposed a machine learning approach to predict adverse drug reaction (ADR) outcomes for given patient characteristics and drug usage [13]. Another study highlights importance of predicting the likelihood of clinical trial side effects using human genetic studies of drug-targeted proteins [14]. From a pharmacogenomics perspective, predicting drug-target interactions using pharmacological similarities of drugs and the US Food and Drug Administration (FDA) Adverse Event Reporting System (FAERS [15]) can be beneficial for drug repositioning and repurposing [16].

FAERS is a voluntary, post-marketing pharmacovigilance tool that can be used to monitor the clinical performance of drugs. In this study, we explore an alternative use of FAERS data to predict compound safety using Medical Dictionary for Regulatory Activities (MedDRA® [17]) terms, which we envision to be useful for future preclinical studies. Our machine learning approach is different from the aforementioned approaches because we not only predict adverse drug

reaction occurences of drugs but most importantly also extract biologically meaningful target-ADR links using *in vitro* secondary pharmacology assay data of more than 2,000 marketed or withdrawn drugs and 184 target assays. Our random forest (RF) model predicts 221 target-ADR associations. We validate the model predictions with canonical target-ADR associations, such as hERG binding causing cardiac arrhythmias. We also encountered unexpected associations which warrant further investigations, such as a link between Phosphodiesterase 3 (PDE3) and several ADRs, including congenital renal and urinary tract disorders. We conclude our study with potential targets that are associated with cardiovascular and renal ADRs to demonstrate the utility and possible impact of this method in drug development and preclinical safety sciences by enabling prediction of ADRs from *in vitro* pharmacological profiles.

Results

Systematic *in vitro* pharmacology of marketed and withdrawn drugs

To link gene targets to ADR occurrence, we utilized *in vitro* pharmacology assay data for 2134 marketed or withdrawn drugs, generated by Novartis, and ADR reports from FAERS (Figure 1A, Supplementary Table 1). Withdrawn drugs and their assay data are also included due to the fact that they are associated with a plethora of ADRs, and thereby constitute an important resource for our predictive approach. Figure 1B summarizes the top 50% of frequently occurring primary indications, classified by the Anatomical Therapeutic Chemical (ATC) codes, of the 2134 drugs using a word cloud. The categories that have the highest number of compounds are antibacterial, ophthalmological, and antineoplastic drugs. The *in vitro* pharmacology assay data includes AC_{50} values for each drug at up to 218 different assays for 184 gene targets (see Supplementary Table 2 for a list of target assays). There are 6 classes of these 184 gene targets, with the majority (47%) of targets falling into G protein-coupled receptors (GPCRs) (Figure 1C). Figure 1D is a heatmap visualization of the *in vitro* pharmacology assay data, where each row is a drug, grouped by their ATC anatomical main group terms [18], and each column is a target assay, grouped by target class. It consists of AC_{50} values of drugs for target assays. The heatmap is not a complete data matrix; 70% of drug-assay combinations have not been tested, i.e. these combinations have NA value for AC_{50}. Nevertheless, our data indicate relatively uniform assaying with respect to the different drug classes.

Analysis of adverse event reports from FAERS connects drugs with human ADRs

We queried FAERS [15] using openFDA [19] for 2134 marketed or withdrawn drugs in November 2018 (FAERS Q4_2018 version) and retrieved 671143 adverse event reports (AERs) using our data extraction criteria (Figure 2A). We only included reports which were submitted by physicians and were annotated as the primary suspect drug [20]. There are 464 drugs that did not have a matching name in FAERS, 341 drugs that did not have any adverse event reports, and 1329 drugs with at least 1 adverse event report. We developed a significance test based on a binomial null distribution and false discovery rate (FDR) multiple testing correction to determine if the observed ADR occurrence was significantly high to be classified as an association (or alternatively no association) between ADR and drug (see Methods for detail). Overall, we observe a positive trend between the number of adverse event reports and the number of ADR associations (Figure 2B). Antineoplastic and immunomodulatory drugs (Figure 2B, blue, N=155) have many ADR associations while the extent of ADR association for antihypertensive drugs (Figure 2B, red, N=35) varies more widely. As an example, we visualized our drug-ADR associations (Figure 2C), in which ADRs are grouped by MedDRA System Organ Class (SOC) level terms and drugs are grouped by ATC anatomical main group terms [18], revealing that ADRs are widespread across organs caused by antineoplastic and immunomodulating agents (Figure 2C, label L), as well as nervous system drugs (Figure 2C, label N).

Random forest model learns relationship between *in vitro* pharmacology and reported ADRs in humans

We developed a machine learning approach to predict ADRs for a given drug from their *in vitro* secondary pharmacology profiles (Figure 3A). We consider this a multi-label classification problem because a given drug can cause multiple ADRs based on its possible engagement with multiple targets and because a single target may be associated with multiple ADRs. We discretized and one-hot encoded our *in vitro* pharmacology assay data (AC_{50} values) into 3 classes: highly active ($AC_{50} < 3$ μM), active (3 μM $\leq AC_{50} \leq 30$ μM) and inactive ($AC_{50} > 30$ μM). Furthermore, we used the binarized drug-ADR associations as described above. We constructed a random forest binary classifier for each ADR under the assumption that ADRs are independent from each other. 413 features were used to predict 321 High Level Group Term (HLGT) ADRs and 21 System Organ Class (SOC) ADRs. The unifying random forest model (binary relevance model), which consists of 321 random forest ADR models, was first trained using 5 fold cross validation where each fold is selected sequentially. We used 1329 drugs for

model construction because these drugs are associated with at least 1 recorded ADR in FAERS. The remaining 805 drugs were not used for training or cross-validation. The model predictions are in probability format, which is used later for target-ADR predictions, and in boolean format (Figure 3A), to enable assessment of model performance via accuracy, macro-precision, macro-recall and Matthew's correlation coefficient (MCC), a performance measure that takes class imbalance into account (Figure 3B). The unifying random forest model performance of SOC ADRs and HLGT ADRs using the full training set (1329 drugs) and the 5 fold cross validation sets (266 drugs, averaged) are depicted in Figure 3B. Accuracy ranges from 0.81 to 0.94, macro-precision ranges from 0.39 to 0.98, macro-recall ranges from 0.22 to 0.63, and MCC ranges from 0.37 to 0.82. Compared to SOC level (21 ADR terms), the finer grain HLGT level (321 ADR terms) had fewer positive drug-ADR associations; additionally, the performance of the HLGT and SOC models are comparable. We therefore proceeded with the HLGT level models for further investigation.

For 55 of the 321 HLGT ADRs, the corresponding random forest models simply predicted zero for all drugs as mostly none (and at most 4) of the 1329 drugs with adverse event reports were associated with those ADRs (Supplemental Table 3). Since these models were not predictive, we did not consider them for further analyses. For the remaining 266 ADRs, we could determine performance metrics (Fig 3C). Accuracy and precision were high, ranging between 0.9 and 1, whilst the recall and MCC range more widely (Fig 3C). This variability occurs for ADRs that have only a few drugs associated with them (Fig 3D). As the number of associated drugs increases and as `more positive training data accumulates, the models learn to better distinguish true positives from false negatives so that their recall and MCC values increase (Fig 3D).

Random forest model predicts expected ADR profiles for anti-hypertensive drugs

As one demonstration of model validation, we analyzed the ADR profiles of 6 subclasses of antihypertensive drugs: adrenergic alpha, adrenergic beta, ACE inhibitors, angiotensin AT2 inhibitors, calcium channel blockers and diuretics (Supplementary Table 4). The signature of the anti-hypertensive drug subclass represents a set of ADRs that were common to all drugs in this subclass. Each antihypertensive drug subclass has a unique ADR fingerprint in the FAERS database which was closely predicted by our random forest model (Figure 3E). The accuracy ranged from 0.984 to 1, with perfect specificity and precision (Supplementary Table 5). The sensitivity ranged from 0.882 to 1, except for the diuretics sub-class, which had a sensitivity of 0.167. This may be because diuretics target the kidney, and not the cardiovascular system as

the rest of the anti-hypertensive drugs do. Of note, the adrenergic alpha and adrenergic beta receptor subclasses maintain distinct profiles in the predicted data. Specifically, the model correctly predicts that adrenergic alpha receptor drugs are associated with suicidal and self injurious behaviors, which has been reported in the literature [21,22].

Random forest model is able to capture ADRs for a test set of drugs

We utilized our random forest model to predict ADRs for 805 drugs that were not used for model construction as they did not have any available ADRs in the FAERS Q4_2018 version, either because there was no match with the drug name or there was no ADR report for that drug submitted to FAERS. We re-queried FAERS for these 805 drugs in the beginning of 2019 (FAERS Q1_2019 version), and retrieved ADRs for 51 of them. These were used as an independent test set to compare their ADR predictions with the actual ADR reports in FAERS Q1_2019. 51 drugs had at least one ADR association in the later version (Supplementary Table 6), and for 18 of them the model made at least one True Positive prediction (TP > 0). The model performance metrics for these 18 drugs range from 0.77 to 0.99 for accuracy, from 0.1 to 0.57 for MCC, from 0.07 to 1 for precision, and from 0.01 to 1 for recall (Figure 3F). The model is able to capture ADRs reported in the latest version of FAERS (Q1_2019) as well as ADRs that are not reported in FAERS database yet but are reported elsewhere. For instance, Glibenclamide (Glyburide), used for treating type 2 diabetes, has 3 TP hits, i.e. the model correctly predicts these 3 ADRs that are reported in the latest version of FAERS. The model suggests 3 more ADR association (FP = 3) with Glibenclamide: (1) epidermal and dermal conditions; (2) gastrointestinal signs and symptoms; (3) respiratory disorders. These ADRs are supported by the Drugs.com database [23], even though they are not reported in the FAERS database. These effects include hives, itching, skin rash, abdominal pain, nausea, shortness of breath and wheezing. Another example, Arformoterol (Aformoterol), is a bronchodilator for chronic obstructive pulmonary disease (COPD) treatment. Our model is able to capture 4 reported ADRs and 3 unreported ADRs including cardiac arrhythmias. According to a safety and efficacy study on Arformoterol, some patients experienced cardiac arrhythmia-related adverse events [25]. Overall, our random forest model proves to be a powerful tool to predict drug-ADR associations from *in vitro* pharmacological drug profiles.

Random forest model predicts 221 target-ADR associations

We used the trained ADR models to predict which target genes are associated with ADRs. To achieve this, we utilized the Gini importance score to rank features for their importance in random forest model performance (Figure 4A). For a given ADR, we selected assays that had multiple AC_{50} features represented in the top 5% of Gini scores ranking (see Methods for detail). We then generated ADR probability predictions for an *in silico* compound that is assumed to target only the selected assay with an AC_{50} value corresponding to a represented feature. We also assumed no available data for all other assays. Using this *in silico* AC_{50} profile as an input to the ADR model, we could generate the ADR probability. By assessing differences in ADR probabilities (two sample t-test, FDR corrected p-value < 0.1) between different AC_{50} classes, e.g. 0-3 μM vs inactive, we predict positive or negative correlations, collectively termed associations, between the selected target assay and ADR. Unsurprisingly, some ADRs did not generate any target associations.

To find biologically meaningful associations, we first filtered out HLGT terms belonging to SOC classes that are not specific to human body parts or only procedural or intervention related (see Methods for detail). Secondly, we filtered out terms that fall under the SOC class neoplasms, since genes are often severely misregulated in cancers and therefore not representative of neoplasm-related ADRs in the organ where the tumor resides. After filtering, we found 221 statistically significant target-ADR associations (Figure 4B, full details including p-values in Supplementary Table 7). 51 out of 184 target assays and 132 out of 321 ADRs are represented (Figure 4B).

To gain confidence in our ADR-target identification procedure, we focused on the $GABA_A$ receptor, the primary target of benzodiazepines (BZD), a drug class known to be psychoactive. Consistently, our model predicts that this ligand-gated chloride ion channel assay is associated with 14 ADRs, 13 of which are neurologically and psychiatrically related, including disturbances in thinking and perception, sleep disorders, depression and suicidal behaviors (Figure 4B, Supplementary Table 7).

Evidence for targets that are predicted to cause cardiovascular-related ADRs

To further validate our model's ability to predict target-ADR associations, we investigated a group of cardiovascular ADRs. We found that hERG binding was associated with cardiac arrhythmias and heart failure at the HLGT level. hERG encodes for a subunit of the cardiac potassium ion channel and contributes to cardiac electrical activity, which is necessary to

regulate the heartbeat. The mechanism of action for drug-induced arrhythmias by blocking hERG has been described in numerous human [26] and animal studies [27], as well as structural modeling [28] studies (Table 1). We did not find an ADR probability associated with 0-3 μM AC_{50} of hERG binding, likely because such strong binding to hERG is a common reason for deprioritizing drug candidates in development [29].

The model predictions also suggest that PDE3 inhibition is associated with cardiac valve disorders (Table 1). PDE3 inhibition is used clinically to treat dilated cardiomyopathy [30], which encapsulates valvular heart disorder. However, the PDE3 therapeutic window is narrow, partially due to complex signaling networks [31], and careful dosing is required to avoid increased mortality in response to treatment.

Our model also predicts that adenosine transporter (AdT) inhibition increases the risk of pericardial disorders (Table 1). For this scenario, we did not find supporting evidence in the literature. The data also suggest that glucocorticoid receptor (GR) binding is more likely to lead to myocardial disorders if the drug has high affinity for GR. This is supported by the finding that glucocorticoid treatment of patients with rheumatoid arthritis increased the risk of myocardial infarction [32]. Furthermore, it is known that dysregulation of glucocorticoids can give rise to cardiotoxicity [33].

Taken together, this investigation of genes associated with cardiovascular ADRs confirms the well-known association of hERG with cardiac arrhythmia, and also highlights ADR associations that would merit further experimental investigation.

Evidence for predicted associations of COX-2, PDE3 and hERG with kidney related ADRs

Another important class of ADRs are related to the kidney (Figure 4B, label: renal). We found COX-2 associated with nephropathies (Table 2), which has been well recognized and evidenced previously [34–36]. Interestingly, another model prediction is PDE3 sensitivity correlating with congenital renal and urinary tract disorders (Table 2). According to a mouse model study [37], PDE3 inhibition could be a contributing factor in Polycystic Kidney Disease (PKD), as PDE3 protein levels are already lower in PKD than WT kidneys. Lastly, we found an unexpected association between hERG and renal disorders (excluding nephropathy) (Table 2). One study has found a loss of hERG function in renal cell carcinoma [38]. In humans, hERG expression in the kidney is much lower than in the heart [39]. Therefore, we conclude that a link between hERG and renal disorders remains a prediction that warrants further investigation.

PDE3 and nuclear hormone receptors AR, ERa and PR are overrepresented in ADR associations

To investigate if the number of different drugs tested for a target assay is predictive to the number of ADRs associated with that target (Figure 4C), we calculated their Pearson correlation coefficient to be moderate (r=0.4). However, some targets had considerably more associated ADRs than other targets that were tested a similar number of times, indicating that more frequently performed assays do not necessarily result in a higher number of associated ADRs (Figure 4C). Out of all target assays, PDE3 was associated with the most ADRs (40) (Figure 4C), falling in a wide range of SOC classes (Figure 4B, Supplementary Table 7). Furthermore, nuclear hormone receptors for androgen (AR), estrogen (ERa) and progesterone (PR) binding assays also have disproportionately many ADR associations, compared to their frequency of testing (Figure 4C). As expected, AR (7/14 ADRs), ERa (9/10 ADRs) but not PR (0/17 ADRs) are associated with sexual reproductive organ and pregnancy-related ADRs (Figure 4B, Supplementary Table 7). Androgen is produced in the adrenal gland [40] and we predict a link between AR with adrenal gland disorders, with evidence in mouse studies [41]. Interestingly, the model predicted 6 ocular ADRs associated to PR, including vision disorders, anterior eye structural change (deposit and degeneration), infections, irritations and inflammations and structural changes (Figure 4B, Supplementary Table 7), for which we could find supporting evidence [42].

Bile salt export pump BSEP associations with ADRs in various organs

BSEP, encoded by *ABCB11* and a member of the superfamily of ATP-binding cassette (ABC) transporters, is most highly expressed in the liver [39]. Drugs that target BSEP are often associated with hepatotoxicity [43]. However, initially, we did not find a BSEP association with hepatic and hepatobiliary disorders. To investigate this false negative prediction, we noted that the dynamic range of the BSEP assay specifically extends up to 300 μM because the liver is likely to experience a higher concentration of orally delivered drugs than the general circulation [44]. As a result, most of our data falls into the ‘inactive’ (>30 μM) class. Consistently, the BSEP inactive feature had the highest Gini score for this HLGT term, while its two active features had much lower Gini scores, falling outside of the top 5%. To take the extended dynamic range into account, we altered the BSEP assay class boundaries to 0-30 μM, 30-300 μM and >300 μM and

retrained the random forest model. In this case, we did find BSEP associated with hepatic and hepatobiliary disorders (Table 3), according to our association criteria (Figure 4A). We repeated this procedure whilst replacing the first class boundary (30 μM) with 100 μM and found the same association again, indicating the robustness of our results. Interestingly, with our original AC_{50} discretizations (Figure 1D), we found BSEP associated with 7 other ADRs from various organ classes (Table 3), much more than other targets that were assayed at a similar frequency (Figure 4C). This suggests that compounds potent against BSEP (AC_{50} < 30 μM) could cause adverse effects in addition to hepatotoxicity, which already occurs at lower potency. We found BSEP associated with urolithiasis and with disorders of the thyroid gland, upper respiratory tract disorders (excl infections), lipid metabolism and central nervous system (Table 3). Since BSEP expression is much lower in these organs [39], we searched the literature for evidence including a substrate of BSEP, bile acid. We could find previous studies linking bile acid to these disorders (Table 3), which suggests an indirect relation between BSEP and these ADRs through bile acid metabolism. Lastly, we found BSEP associated with foetal complications and pregnancy conditions (Table 3), both supported through prior studies that link BSEP transient neonatal cholestasis and intrahepatic cholestasis of pregnancy, respectively [45,46].

Discussion

In this study we have taken a machine learning approach to predict human ADRs from the *in vitro* secondary pharmacology profiles of a large number of marketed and withdrawn drugs. Several prior studies focus on predicting ADRs directly from chemical drug structure [47,48]. However, utilizing functional information such as *in vitro* pharmacological targeting of common (off) targets represents a viable alternative to bridge the complex relationship between drugs and their effects in the human body [4].

Our random forest model performance metrics are good considering the sparse coverage (2134 drugs) over a large input space (3^{184} possibilities) and partial overlap with ADR reporting for these drugs, making ADR occurrence prediction effectively a one shot learning task. Importantly, optimizing test performance was not the main objective of this study. Instead, we endeavored to find biologically meaningful target-ADR associations. To achieve this without relying on test performance, we trained on all data and made use of Gini scores to robustly select relevant features for ADR probability predictions. Our novel method for target-ADR

associations was able to recapitulate well recognized causal relations, such as hERG with cardiac arrhythmias. For others, we were able to find literature evidence in animal or *in vitro* studies but our study is, to our knowledge, a first in human report. Another fraction of target-ADR associations represents predictions of novel, unexpected or little known associations, such as Adenosine Transporter (AdT) and pericardial disorders, for which we could find little evidence other than our analysis of adverse event reports. Overall, our 221 associations form a rich resource that can be used for further mechanistic studies in the drug discovery process.

Our random forest model is agnostic to molecular mechanisms; therefore, resulting associations could arise from indirect regulation. A likely example is the bile transporter BSEP, which is associated with numerous ADRs, although it is most highly expressed in the liver and kidney. We have related our findings to evidence that misregulation of its substrate, bile acid, could result in disorders related to kidney stones, lipid metabolism, thyroid gland, respiratory system and central nervous system. This also indicates the strength of our approach, which can relate genes to physiological processes unbiasedly in humans, without any interventions or large scale genome-wide association studies, but solely with voluntary adverse event reporting.

While we recommend this approach to find target-ADR associations to impact safety awareness in drug discovery, we are also aware of the limitations. Firstly, the presented analyses are limited by the input data. The *in vitro* data matrix is incomplete (targets in the *in vitro* pharmacology panel cover a small fraction of the biological target space and not all drugs were tested in all assays). We recognize that the present set of targets is biased towards the GPCR target family with limited representation of other therapeutic or ADR-associated targets such as ion channels and kinases. Also, data are influenced by prior knowledge; for example, more than 87% of all drugs in the set were tested for hERG activity. High affinity (lower AC_{50} value) for hERG is associated with higher probability for QT prolongation for human and non-human preclinical species [26,27]. As discussed earlier, there are not many drugs with a hERG AC_{50} value in the highly active class (0-3 μM), which is a commonly encountered roadblock for drug candidates to progress towards clinical trials [29]. Only about 10% of all drugs fall into the highly active class in our assay data. To limit feature engineering, our AC_{50} discretization into three classes (Figure 1D) was kept uniform across all assays. Notably for the BSEP assay, the dynamic range extends up to 300 μM and as a result most of our data falls into the ‘inactive’ (>30 μM) class. Consequently, we did not find the expected association with hepatotoxicity. We

rectified this by reclassifying the BSEP assay data according to levels required for hepatotoxicity of BSEP inhibition [49,50] and indeed recovered the expected association.

Secondly, *in vitro* potency is a very simplified marker of clinical effect, and does not take into account prolonged dosing, comorbidity or pharmacokinetic/pharmacodynamic relationships (e.g. therapeutic window). For some targets, although very rarely, non-human proteins were assayed (e.g. rat brain was used as a source for the benzodiazepine receptor) which may not be a direct correlate of the human protein. Further development of the model would require addition of parameters on occupancy and pharmacodynamic components for more precision and enhanced predictive value.

Thirdly, the FAERS database is biased. For example, drug-ADR associations may be mislabeled, e.g. anti-hypertensives are often reported as associated with hypertension as an ADR, rather than as the indication. This and other limitations are discussed by Maciejewski et al. [20] with suggestions and methodology for further refinement for the method. Additionally, the FAERS database does not contain information on the total number of patients exposed to a particular drug, nor is it necessarily a reflection of the true incidence or frequency of ADRs.

Our model assumption that ADRs are independent from each other is a reasonable first step to tackle our scientific challenge. However, some ADRs can be the prerequisite for others (e.g. hypotension leading to reflex tachycardia). We leave a model extension to incorporate these dependencies as future work. For target-ADR associations, we utilized our random forest model for a single drug at a time. One can repurpose our model to predict possible ADRs from combination drug therapies and likelihood of drug-drug interactions. In principle, this can be extended straightforwardly for combination therapies by merging the *in vitro* data from the individual compounds. Offside and Twosides databases can be used for validation [51]. Similarly, our model can be utilized for drug repositioning and repurposing, using similar target-ADR profiles. We leave these endeavors as future work.

In conclusion, our target ADR associations provide a comprehensive resource to support drug development and future human biology studies.

Methods

In vitro secondary pharmacology assays for marketed drugs

AC_{50} values of 2134 marketed drugs (Supplementary Table 1) were measured in up to 218 different *in vitro* secondary pharmacology assays. Compounds were obtained from the Novartis Institutes of Biomedical Research (NIBR) compound library and tested in a panel of *in vitro* biochemical and cell-based assays at Eurofins and at NIBR in concentration-response (8 concentrations, half-log dilutions starting at 30 µM). Assay formats varied from radioligand binding to isolated protein to cellular assays. Example protocols may be found at https://www.eurofinsdiscoveryservices.com/cms/cms-content/services/in-vitro-assays/ . Normalized concentration response curves were fitted using a four parameter logistic equation with internally developed software (Helios). The equation used is for a one site sigmoidal dose response curve Y as a function of tested concentrations X: $Y(X)=A+(B-A)/(1+(X/C)^{D})$, with fitted parameters A=min(Y), B=max(Y), $C=AC_{50}$ and exponent D. By default, A is fixed at 0, whereas B is not fixed.

If a drug was not tested against a specific assay, the AC_{50} value was set to NA (not available). AC_{50} values from similar assays with the same gene target were merged to reduce the NA data and features in the random forest model; this procedure resulted in 184 different target assays (Supplementary Table 2). In case any merged assays had multiple AC_{50} values for the same drug, we averaged these geometrically to take into account variation over orders of magnitudes. In figures 1D and 2C, the drugs are classified according to their annotated Anatomical Therapeutic Chemical (ATC) code [18]. In case of multiple ATC codes, we assigned the most frequent level 1 code.

Mining adverse event reports of marketed drugs using OpenFDA

In this study, we utilized openFDA to acquire FAERS reports related to the query compounds [15,19]. This Elasticsearch-based API provides a raw download access to a large volume of structured datasets, including adverse events reports from FAERS.

We used generic compound names (e.g. “Amoxicillin”) to query through the openFDA interface, accessed programmatically using Python. In order to maximize the coverage over FDA datasets, we normalized generic names to uppercase format followed by a name similarity metric to filter out unrelated records in our analysis. We included reports when the Jaro similarity between the query generic name and reported compound name was equal or greater than 0.8.

As the FAERS database contains information voluntarily submitted by healthcare professionals, consumers, lawyers and manufacturers, adverse event reports may be duplicated by multiple parties per event, and may be more likely to contain incorrect information if submitted by a non-medical professional. To reduce reporting bias and increase report information accuracy, we only analyzed reports submitted by physicians (data field: 'qualification' = 1). In this subset of adverse event reports, the data were further filtered by reported drug characterization, which indicates how the physician characterized the role of the drug in the patient's adverse event. A drug can be characterized as a primary suspect drug, holding a primary role in the cause of the adverse event (data field: 'drugcharacterization' = 1); a concomitant drug ('drugcharacterization' = 2); or an interacting drug ('drugcharacterization' = 3). Here, we included only primary suspect drug reports. Without this restriction, model performances did not improve. We obtained all adverse events reports corresponding to the query compound that passed through the aforementioned filters.

Adverse event report descriptions are coded as medical terms of MedDRA terminology [17]. Medical observations can be reported using 5 hierarchical levels of medical terminology, ranging from a very general System Organ Class term (e.g. gastrointestinal disorders) to a very specific Lowest Level Term (e.g. feeling queasy). Each term is linked to only one term on a higher level. For each report, we recorded all MedDRA Reaction terms (data field: "reactionmeddrapt") at the Preferred Term level and mapped these Preferred Terms to Higher Level Group Term and System Organ Class level. For each (ADR term, drug) tuple, we then calculated the ADR occurrence, defined as the following fraction: number of adverse event reports containing that ADR term relative to the total number of ADR reports for that drug.

Random forest models

To construct and train our models (Figure 3A), we used AC_{50} values for a panel of target assays for marketed drugs (model input; independent variable) and ADR occurrences of the compounds (model output/predictions; dependent variable). Since there may be several ADRs associated with any given drug, we considered this a multi-label learning problem [52]. We took a "first-order strategy", i.e. we assume there is no correlation between different ADRs, and a "divide and conquer" strategy, i.e. we decompose our multi-label learning task into n independent binary classification problems, where n is the number of different ADR terms in our output data (n = 26 for SOC and n = 321 for HLGT level respectively). We built a random forest

[53] binary classifier for each ADR using Binary Relevance with the random forest modeling option in mldr package [54] and utiml package in R [55].

To define the features for the random forest models, we discretized and one-hot encoded our input AC_{50} values. We defined 3 classes (levels) of AC_{50} ranges for each target assay.

- Highly active class: AC_{50} in [0, 3 μM)
- Active class: AC_{50} in [3 μM, 30 μM]
- Inactive class: AC_{50} greater than 30 μM

If the AC_{50} value is NA, the values for all Classes are 0. Each drug has AC_{50} values for 184 (merged) assays, so there are 184x3 = 552 binary features to represent our input data. Features consisting of only 0 values were removed, resulting in 413 input features used for model construction.

The observed ADR occurrences were discretized into binary dependent variables. To achieve this, first let N_d be the total number of ADR reports for a given drug. The probability to observe an ADR occurrence $O^{ADR} = X / N_d$ at random is equivalent to choosing that ADR X times out of N_d with X distributed binomially: $X \sim bin(N_d, p=1/n)$. Here, n represents the total number of ADRs as defined above. Under this null distribution, we calculate the p-values for all observed ADR occurrences O^{ADR} for a given drug, and then perform a Benjamini-Hochberg False Discovery Rate (FDR) correction (using the Python statsmodels package). If an FDR-corrected p-value is < 0.01, then the ADR value for that drug is 1, reflecting an association; 0 otherwise.

All random forest models were first trained using 5 fold cross validation and each fold is selected sequentially. 1063 drugs were used for training and 266 drugs were used for testing in each fold. Then, the drugs with at least 1 ADR report are used as a training set. For a given (drug) input of AC_{50} values and ADR, the random forest model output, termed ADR probability, can be interpreted as the probability that the ADR is associated with the drug. To enable direct comparison of model predictions with binarized ADR occurrences, we binarized these ADR probabilities with a simple threshold value of 0.5. These binary values were used for training, cross validation and to calculate classification performance metrics (Figure 3B,C).

We evaluated our models based on four metrics: accuracy, Matthew's correlation coefficient (MCC), macro-precision and macro-recall. These metrics are calculated using utiml, mldr and mltools packages in R [54,55] (https://github.com/ben519/mltools). The corresponding metrics for

the each ADR model are accuracy, precision and recall, which are calculated using mltools package in R for MCC, and the definitions:

- Accuracy = $(TP + TN)\ /\ (TP + TN + FP + FN)$
- Precision = $TP\ /\ (TP + FP)$
- Recall = $TP\ /\ (TP + FN)$
- MCC = $(TP * TN - FP * FN)\ /\ SQRT\ ((TP + FP) * (TP + FN) * (TN + FP) * (TN + FN))$

Determination of target-ADR associations

To find associations between gene target assays and ADRs (Figure 4), we first generated ADR probabilities specific to a given assay. As a model input, one out of its three random forest input features' value was set to 1 and all others to 0. This simulates the scenario of an *in silico* compound that is potent with an AC_{50} value in the range corresponding to the positive feature only. We then utilized the ADR's random forest model, pre-trained on all available marketed drug data (see previous section), to calculate the resulting ADR probability. We repeated this procedure for each feature of all assays and each ADR.

To select the predictive features for a given ADR, we ordered the pre-trained random forest model's input features according to their Gini importance score [56] and denote the top 5% as significant features. Our criteria for a gene (target assay) - ADR pair were:

- For a given ADR: at least 2 out of 3 assay features need to be significant in order to make a reliable comparison between the ADR probabilities with respect to AC_{50} values.
- At least one of the ADR probabilities of the significant features has to be larger than zero.

We filtered out target-ADR pairs if the ADR term maps to the following SOC classes, which are not specific to body parts or underlying human biology:

- general disorders and administration site conditions
- injury, poisoning and procedural complications
- investigations
- neoplasms benign, malignant and unspecified (incl cysts and polyps)
- poisoning and procedural complications
- social circumstances
- surgical and medical procedures

To ensure the reproducibility of the target-ADR pair selection procedure, we repeated the random forest model training with different seeds for a total of 5 times. We then took the union of the 5 sets of target-ADR pairs and discarded pairs that were only found once out of 5 runs. Finally, to determine if the mean ADR probabilities between the selected AC_{50} classes were statistically significantly different, we performed a two-sample t-test with sample sizes equal to the number of times a class was selected (ranging from 2 to 5 times) using the Python scikit.stats package. In case all three AC_{50} classes were represented, we tested the highly active versus inactive class. We then performed a Benjamini-Hochberg FDR correction. If the FDR-corrected p-value is < 0.1, then the target-ADR pair is considered a statistically significant association (Figure 4B, Supplementary Table 7).

To evaluate the relation between the HLGT level ADR term hepatic and hepatobiliary disorders and target assay BSEP, we also trained and analyzed two random forest models as described above to find target-ADR pairs but with only the BSEP assay data discretized with class boundaries [0, 30 μM), [30, 300 μM] and >300 μM or [0, 100 μM), [100, 300 μM] and >300 μM.

Figure Legends

Figure 1. Major elements of the target-ADR association analysis

A. **Schematic outline of target-ADR pair determinations.** The observed relations (solid lines) between drugs and adverse drug reactions (ADRs) are determined by post-marketing pharmacovigilance and between drugs and their (off) targets by *in vitro* pharmacology. This approach enables prediction of associations (dashed line) between targets and ADRs through random forest modelling.

B. **Representation of drug classes in word cloud.** The cloud displays the top 50% most frequently occurring drug classes, representing 2134 drugs, in the Novartis *in vitro* pharmacology data warehouse. Size of the font of the drug class reflects the number of associated drugs.

C. **Target class distribution in the Novartis *in vitro* secondary pharmacology assay panel.** The 184 targets in the Novartis assay panel cover 6 target classes. Almost half of the target assays belong to the G protein-coupled receptor (GPCR) class.

D. **Novartis target panel potency (AC_{50}) heatmap.** The profile consists of the AC_{50} values of 184 target assays for 2134 drugs. We considered an AC_{50} value less than 3 μM as highly active (red), between 3 μM and 30 μM as active (blue) and greater than 30 μM as inactive (yellow). No data for a drug-target pair is labeled as NA (white). Drugs are grouped (horizontally) by their Anatomical Therapeutic Chemical (ATC) codes. Assays are grouped (vertically) by target class.

FIGURE 1

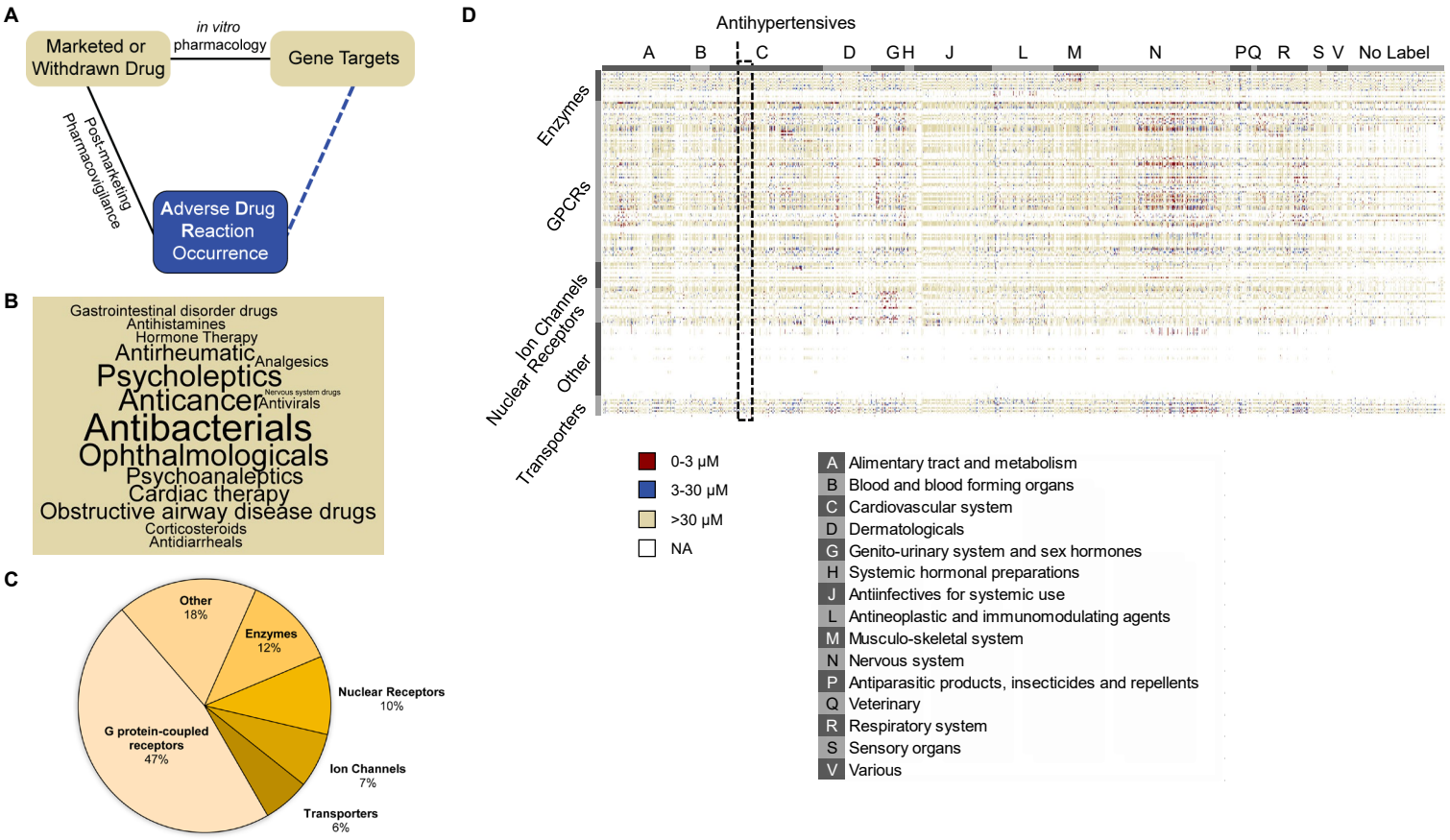
A
Marketed or Withdrawn Drug
in vitro pharmacology
Gene Targets
Post-marketing Pharmacovigilance
Adverse Drug Reaction Occurrence
B
Gastrointestinal disorder drugs
Antihistamines
Hormone Therapy
Antirheumatic
Analgesics
Psycholeptics
Nervous system drugs
Anticancer
Antivirals
Antibacterials
Ophthalmologicals
Psychoanaleptics
Cardiac therapy
Obstructive airway disease drugs
Corticosteroids
Antidiarrheals
C
Other 18%
Enzymes 12%
Nuclear Receptors 10%
Ion Channels 7%
Transporters 6%
G protein-coupled receptors 47%
D
Antihypertensives
A B C D GH J L M N PQ R S V No Label
Enzymes
GPCRs
Ion Channels
Nuclear Receptors
Other
Transporters
0-3 μM
3-30 μM
>30 μM
NA
A Alimentary tract and metabolism
B Blood and blood forming organs
C Cardiovascular system
D Dermatologicals
G Genito-urinary system and sex hormones
H Systemic hormonal preparations
J Antiinfectives for systemic use
L Antineoplastic and immunomodulating agents
M Musculo-skeletal system
N Nervous system
P Antiparasitic products, insecticides and repellents
Q Veterinary
R Respiratory system
S Sensory organs
V Various

Figure 2. Retrieval of Adverse Event Reports from the FDA Adverse Event Reporting System (FAERS) database

A. **Flow chart of the programmatic strategy for Adverse Event Report (AER) retrieval from FAERS by using openFDA.** 'is qualification = 1' is a positive filter for AERs that were reported by physicians. 'is drugcharacterization' = 1 is a positive filter for drugs that are annotated as the primary suspect drug, which hold a primary role in the cause of the adverse event.

B. **Scatter plot of the number of associated ADRs for drugs as a function of the number of AERs retrieved for each drug (N_{drugs} = 1329).** Drugs without any reported ADR are not shown.

C. **Heatmap of ADR profiles (discretized as used for input of random forest model) for all marketed drugs used in this study (N_{drugs} = 2134).** Drugs are clustered (horizontally) according to their ATC drug classes (A-V, or No label if without any ATC code) and HLGT (high level group term) ADRs are grouped (vertically) according to the parent System Organ Class (SOC) level listed in the legend.

FIGURE 2

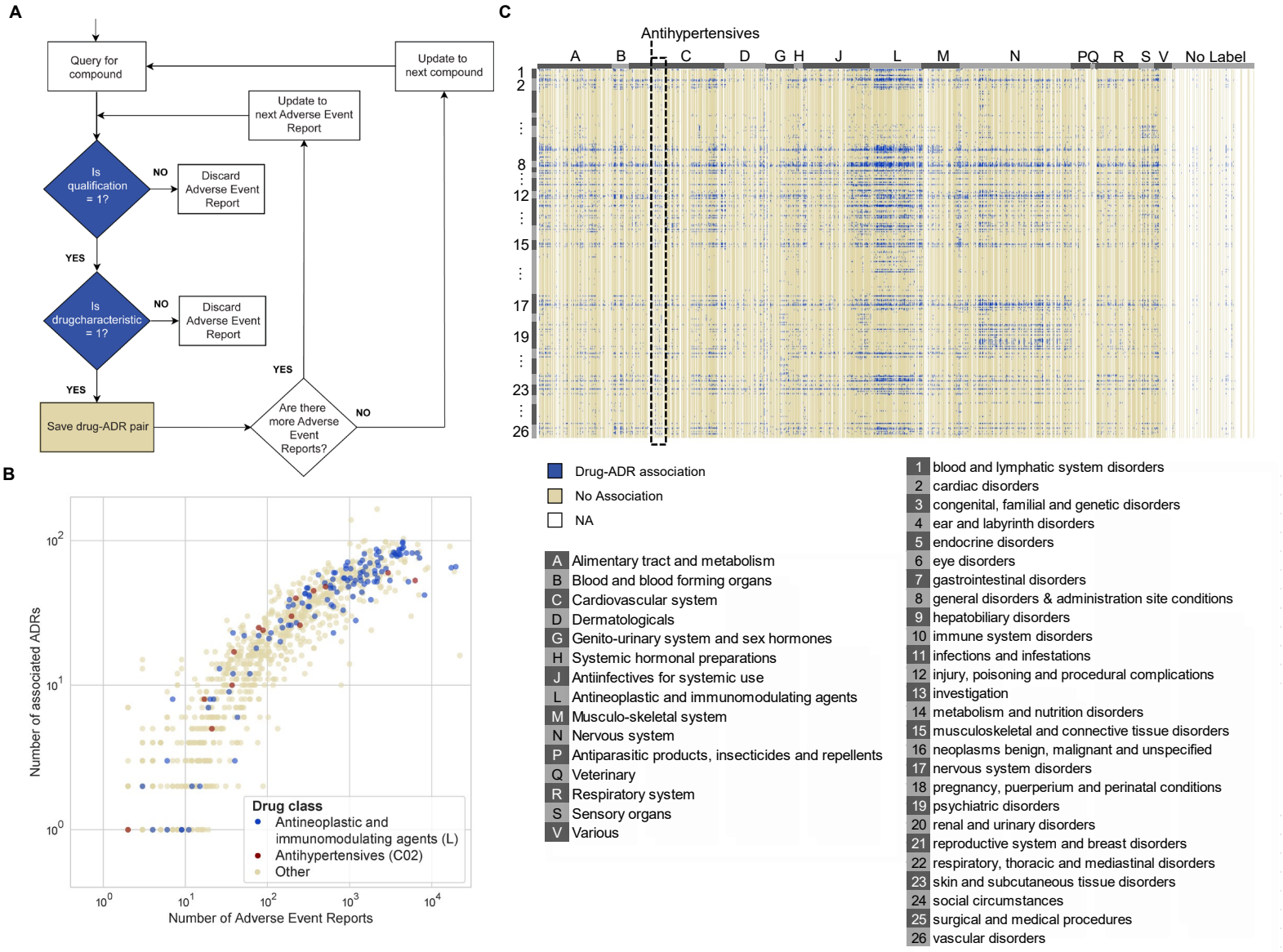
A
Query for compound
Update to next compound
Update to next Adverse Event Report
Is qualification = 1?
NO
Discard Adverse Event Report
YES
Is drugcharacteristic = 1?
NO
Discard Adverse Event Report
YES
Save drug-ADR pair
YES
Are there more Adverse Event Reports?
NO
B
Number of associated ADRs
Number of Adverse Event Reports
Drug class
Antineoplastic and immunomodulating agents (L)
Antihypertensives (C02)
Other
C
Antihypertensives
A B C D G H J L M N PQ R S V No Label
Drug-ADR association
No Association
NA
A Alimentary tract and metabolism
B Blood and blood forming organs
C Cardiovascular system
D Dermatologicals
G Genito-urinary system and sex hormones
H Systemic hormonal preparations
J Antiinfectives for systemic use
L Antineoplastic and immunomodulating agents
M Musculo-skeletal system
N Nervous system
P Antiparasitic products, insecticides and repellents
Q Veterinary
R Respiratory system
S Sensory organs
V Various
1 blood and lymphatic system disorders
2 cardiac disorders
3 congenital, familial and genetic disorders
4 ear and labyrinth disorders
5 endocrine disorders
6 eye disorders
7 gastrointestinal disorders
8 general disorders & administration site conditions
9 hepatobiliary disorders
10 immune system disorders
11 infections and infestations
12 injury, poisoning and procedural complications
13 investigation
14 metabolism and nutrition disorders
15 musculoskeletal and connective tissue disorders
16 neoplasms benign, malignant and unspecified
17 nervous system disorders
18 pregnancy, puerperium and perinatal conditions
19 psychiatric disorders
20 renal and urinary disorders
21 reproductive system and breast disorders
22 respiratory, thoracic and mediastinal disorders
23 skin and subcutaneous tissue disorders
24 social circumstances
25 surgical and medical procedures
26 vascular disorders

Figure 3. Application of the random forest model to characterize drug-ADR associations

A. **Schematic representation of the machine learning approach.** Using input data, which is a discretized AC_{50} *in vitro* pharmacological profile, we built a separate random forest model for each adverse drug reaction (ADR) that predicts the probability of a drug causing that ADR. For training we used all drugs for which we could retrieve FAERS Q4_2018 AERs (N_{drugs} = 1329).

B. **Summary statistics of overall model performance.** We developed two unified random forest models based on two hierarchical levels of organ class specifications. The high level group term (HLGT; blue) unified random forest model consists of 321 ADR random forest models whereas the system organ class (SOC; yellow) unified random forest model consists of 26 ADR random forest models. The performance of the HLGT and SOC models is similar, except in few cases when the HLGT model outperforms the SOC model. (MCC: Matthew's correlation coefficient). Training reflects performance after model training on all drugs (see A). 5-Fold cross validation results are averaged over each fold.

C. **Box plots indicating the distributions of the training performance metrics** (as in B) for all random forest models of each individual HLGT ADR (N_{ADRs} = 266).

D. **Scatter plot of the random forest models' recall (all metrics as in C) as a function of number of associated ADRs**, which served as positive training examples. Colors indicate model precision and circle size reflects the MCC.

E. **ADR predictions for anti-hypertensive drugs with different pharmacological targets**. For a set of 22 antihypertensive drugs, we visualized the association between the drugs and HLGT-level ADRs (left). Using the ADR random forest models, we predicted the differences in ADR associations between antihypertensive drugs representing various pharmacological targets (right; overall 36 of the HLGT terms are visualized). True negative predictions (285 HLGT-level ADRs) were omitted from this visualization.

F. **Test set predictions of ADRs for drugs with AERs only in FAERS Q1_2019** (N_{drugs} = 23), which have an *in vitro* pharmacological profile but did not have AERs reported in FAERS Q4_2018, and were therefore not used for training. However, we could retrieve AERs in FAERS Q1_2019. This enabled us to assess model performances for each HLGT-level ADR. Values in the heatmap reflect the average over those ADRs.

FIGURE 3

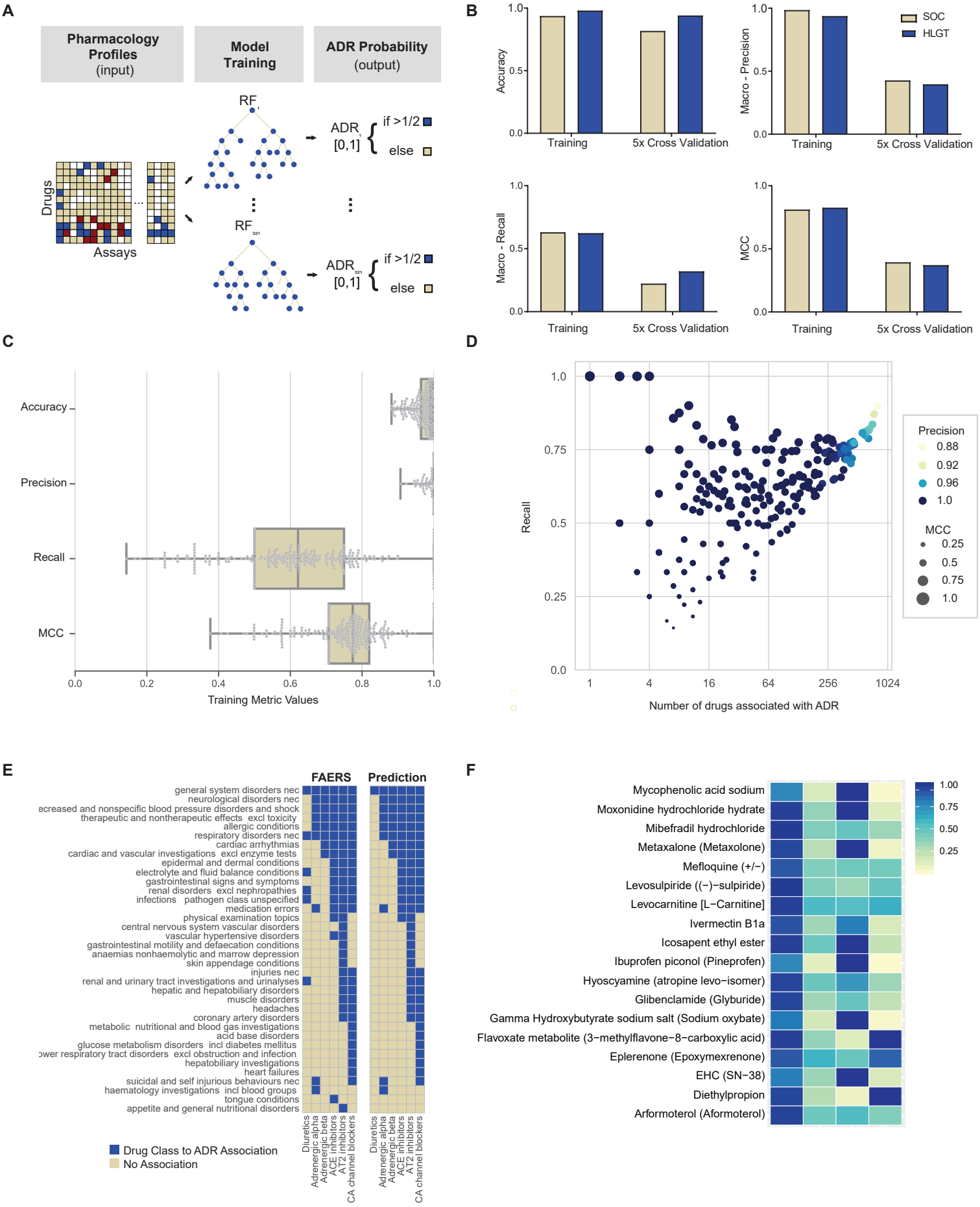
A
Pharmacology Profiles (input)
Model Training
ADR Probability (output)
Drugs
Assays
RF
ADR
[0,1]
if >1/2
else
B
Accuracy
Macro - Precision
Macro - Recall
MCC
Training
5x Cross Validation
SOC
HLGT
C
Accuracy
Precision
Recall
MCC
Training Metric Values
D
Recall
Number of drugs associated with ADR
Precision
0.88
0.92
0.96
1.0
MCC
0.25
0.5
0.75
1.0
E
FAERS
Prediction
general system disorders nec
neurological disorders nec
ecreased and nonspecific blood pressure disorders and shock
therapeutic and nontherapeutic effects excl toxicity
allergic conditions
respiratory disorders nec
cardiac arrhythmias
cardiac and vascular investigations excl enzyme tests
epidermal and dermal conditions
electrolyte and fluid balance conditions
gastrointestinal signs and symptoms
renal disorders excl nephropathies
infections pathogen class unspecified
medication errors
physical examination topics
central nervous system vascular disorders
vascular hypertensive disorders
gastrointestinal motility and defaecation conditions
anaemias nonhaemolytic and marrow depression
skin appendage conditions
injuries nec
renal and urinary tract investigations and urinalyses
hepatic and hepatobiliary disorders
muscle disorders
headaches
coronary artery disorders
metabolic nutritional and blood gas investigations
acid base disorders
glucose metabolism disorders incl diabetes mellitus
ower respiratory tract disorders excl obstruction and infection
hepatobiliary investigations
heart failures
suicidal and self injurious behaviours nec
haematology investigations incl blood groups
tongue conditions
appetite and general nutritional disorders
Diuretics
Adrenergic alpha
Adrenergic beta
ACE inhibitors
AT2 inhibitors
CA channel blockers
Drug Class to ADR Association
No Association
F
Mycophenolic acid sodium
Moxonidine hydrochloride hydrate
Mibefradil hydrochloride
Metaxalone (Metaxolone)
Mefloquine (+/−)
Levosulpiride ((−)−sulpiride)
Levocarnitine [L−Carnitine]
Ivermectin B1a
Icosapent ethyl ester
Ibuprofen piconol (Pineprofen)
Hyoscyamine (atropine levo−isomer)
Glibenclamide (Glyburide)
Gamma Hydroxybutyrate sodium salt (Sodium oxybate)
Flavoxate metabolite (3−methylflavone−8−carboxylic acid)
Eplerenone (Epoxymexrenone)
EHC (SN−38)
Diethylpropion
Arformoterol (Aformoterol)

Figure 4. Random forest model predicts target-ADR associations

A. **Schematic outline of the *in silico* ADR-target predictions**. For an ADR of interest, we determined the top 5% of features from the corresponding trained random forest model, ranked according to their Gini importance scores, which measures their contribution to the predictive power of the model. If at least two features (e.g. as depicted: highly active and inactive) from the same target assay are within that top 5%, we determined the ADR probabilities for the simulated cases where an *in silico* compound would target those assay AC_{50} classes only. The ADR probabilities of those simulated cases can then be compared to determine the concentration dependence of the ADR probability. If there is a non-zero correlation between AC_{50} values and ADR probabilities, we conclude that there is an association between the respective ADR and target. For full detail see the Methods.

B. **Heatmap showing the resulting 221 predicted target-ADR associations (blue)**. Drugs (horizontal) are listed alphabetically, and HLGT ADRs (vertical) are grouped according to their parent SOC level. For a full description of all target-ADR associations and their ADR probabilities, see Supplementary Table 7.

C. **Scatter plot of each target (assay) showing the number of ADR associations as a function of number of assayed drugs.**

FIGURE 4

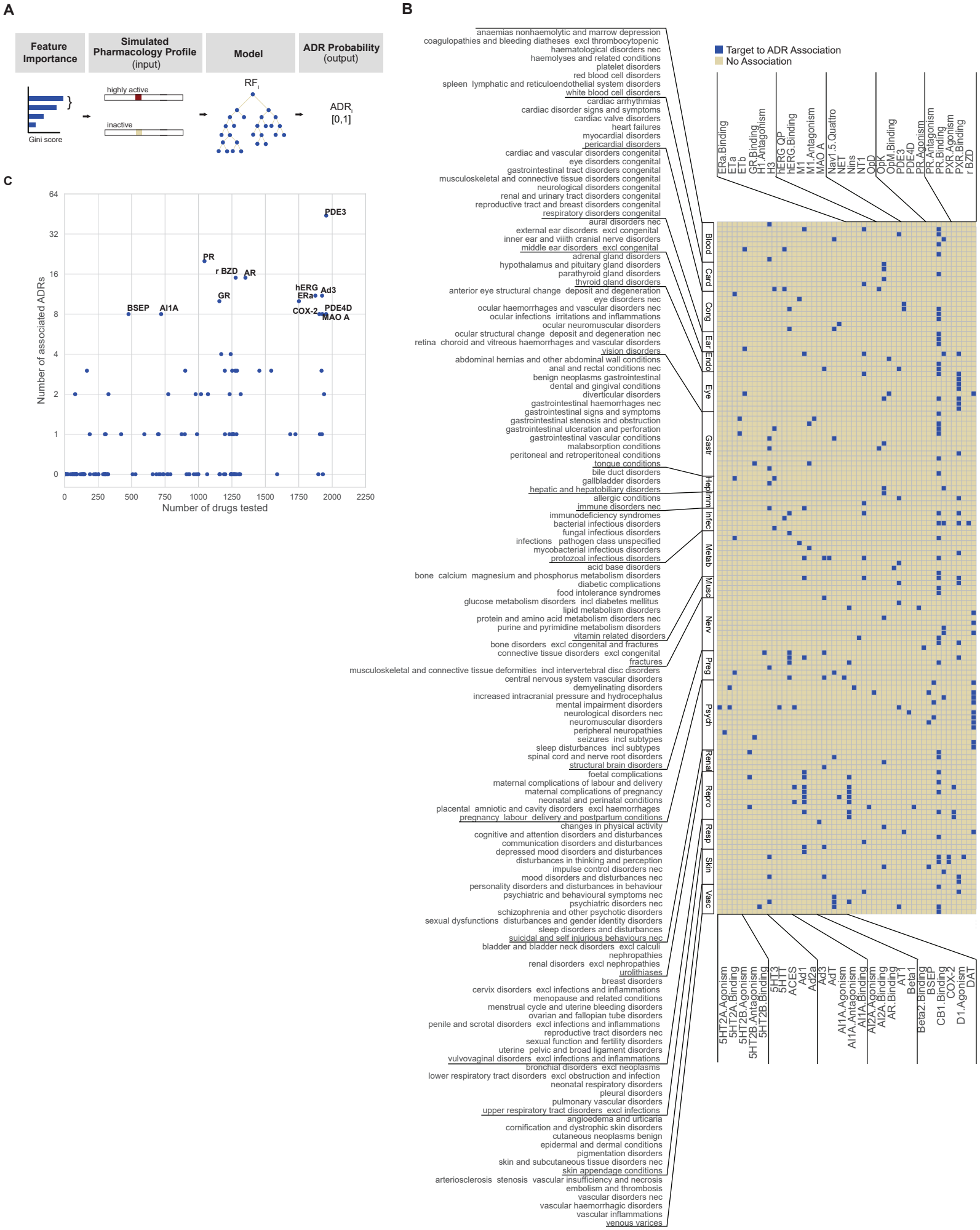
A
Feature Importance
Simulated Pharmacology Profile (input)
Model
ADR Probability (output)
Gini score
highly active
inactive
RF_i
ADR_i [0,1]
B
Target to ADR Association
No Association
ERa.Binding
ETa
ETb
GR.Binding
H1.Antagonism
H3
hERG.QP
hERG.Binding
M1
M1.Antagonism
MAO A
Nav1.5.Quattro
NET
Nins
NT1
OpD
OpK
OpM.Binding
PDE3
PDE4D
PR.Antagonism
PR.Binding
PXR.Agonism
PXR.Binding
r.BZD
Blood
Card
Cong
Ear
Endo
Eye
Gastr
Hepb
Immn
Infec
Metab
Musc
Nerv
Preg
Psych
Renal
Repro
Resp
Skin
Vasc
anaemias nonhaemolytic and marrow depression
coagulopathies and bleeding diatheses excl thrombocytopenic
haematological disorders nec
haemolyses and related conditions
platelet disorders
red blood cell disorders
spleen lymphatic and reticuloendothelial system disorders
white blood cell disorders
cardiac arrhythmias
cardiac disorder signs and symptoms
cardiac valve disorders
heart failures
myocardial disorders
pericardial disorders
cardiac and vascular disorders congenital
eye disorders congenital
gastrointestinal tract disorders congenital
musculoskeletal and connective tissue disorders congenital
neurological disorders congenital
renal and urinary tract disorders congenital
reproductive tract and breast disorders congenital
respiratory disorders congenital
aural disorders nec
external ear disorders excl congenital
inner ear and viiith cranial nerve disorders
middle ear disorders excl congenital
adrenal gland disorders
hypothalamus and pituitary gland disorders
parathyroid gland disorders
thyroid gland disorders
anterior eye structural change deposit and degeneration
eye disorders nec
ocular haemorrhages and vascular disorders nec
ocular infections irritations and inflammations
ocular neuromuscular disorders
ocular structural change deposit and degeneration nec
retina choroid and vitreous haemorrhages and vascular disorders
vision disorders
abdominal hernias and other abdominal wall conditions
anal and rectal conditions nec
benign neoplasms gastrointestinal
dental and gingival conditions
diverticular disorders
gastrointestinal haemorrhages nec
gastrointestinal signs and symptoms
gastrointestinal stenosis and obstruction
gastrointestinal ulceration and perforation
gastrointestinal vascular conditions
malabsorption conditions
peritoneal and retroperitoneal conditions
tongue conditions
bile duct disorders
gallbladder disorders
hepatic and hepatobiliary disorders
allergic conditions
immune disorders nec
immunodeficiency syndromes
bacterial infectious disorders
fungal infectious disorders
infections pathogen class unspecified
mycobacterial infectious disorders
protozoal infectious disorders
acid base disorders
bone calcium magnesium and phosphorus metabolism disorders
diabetic complications
food intolerance syndromes
glucose metabolism disorders incl diabetes mellitus
lipid metabolism disorders
protein and amino acid metabolism disorders nec
purine and pyrimidine metabolism disorders
vitamin related disorders
bone disorders excl congenital and fractures
connective tissue disorders excl congenital
fractures
musculoskeletal and connective tissue deformities incl intervertebral disc disorders
central nervous system vascular disorders
demyelinating disorders
increased intracranial pressure and hydrocephalus
mental impairment disorders
neurological disorders nec
neuromuscular disorders
peripheral neuropathies
seizures incl subtypes
sleep disturbances incl subtypes
spinal cord and nerve root disorders
structural brain disorders
foetal complications
maternal complications of labour and delivery
maternal complications of pregnancy
neonatal and perinatal conditions
placental amniotic and cavity disorders excl haemorrhages
pregnancy labour delivery and postpartum conditions
changes in physical activity
cognitive and attention disorders and disturbances
communication disorders and disturbances
depressed mood disorders and disturbances
disturbances in thinking and perception
impulse control disorders nec
mood disorders and disturbances nec
personality disorders and disturbances in behaviour
psychiatric and behavioural symptoms nec
psychiatric disorders nec
schizophrenia and other psychotic disorders
sexual dysfunctions disturbances and gender identity disorders
sleep disorders and disturbances
suicidal and self injurious behaviours nec
bladder and bladder neck disorders excl calculi
nephropathies
renal disorders excl nephropathies
urolithiases
breast disorders
cervix disorders excl infections and inflammations
menopause and related conditions
menstrual cycle and uterine bleeding disorders
ovarian and fallopian tube disorders
penile and scrotal disorders excl infections and inflammations
reproductive tract disorders nec
sexual function and fertility disorders
uterine pelvic and broad ligament disorders
vulvovaginal disorders excl infections and inflammations
bronchial disorders excl neoplasms
lower respiratory tract disorders excl obstruction and infection
neonatal respiratory disorders
pleural disorders
pulmonary vascular disorders
upper respiratory tract disorders excl infections
angioedema and urticaria
cornification and dystrophic skin disorders
cutaneous neoplasms benign
epidermal and dermal conditions
pigmentation disorders
skin and subcutaneous tissue disorders nec
skin appendage conditions
arteriosclerosis stenosis vascular insufficiency and necrosis
embolism and thrombosis
vascular disorders nec
vascular haemorrhagic disorders
vascular inflammations
venous varices
5HT2A.Agonism
5HT2A.Binding
5HT2B.Agonism
5HT2B.Antagonism
5HT2B.Binding
5HT3
5HTT
ACES
Ad1
Ad2a
Ad3
AdT
Al1A.Agonism
Al1A.Antagonism
Al1A.Binding
Al2A.Agonism
Al2A.Binding
AR.Binding
AT1
Beta1
Beta2.Binding
BSEP
CB1.Binding
COX-2
D1.Agonism
DAT
C
Number of associated ADRs
Number of drugs tested
PDE3
PR
r BZD
AR
hERG
Ad3
ERa
GR
BSEP
Al1A
COX-2
PDE4D
MAO A

Table 1 Predicted associations between targets and cardiac ADRs.

High Level Group Terms (HLGT; MedDRA) associations with targets and Adverse Drug Reaction (ADR) probability in three concentration ranges (third column). Evidence of the ADR-target pairs were obtained from peer reviewed publications (fourth column). hERG: human Ether-a-go-go-Related Gene associated potassium channel; PDE3: phosphodiesterase-3 enzyme; GR: glucocorticoid receptor; AdT: Adenosine transporter; COX-2: cyclooxygenase enzyme, type 2.

Cardiac Disorder HLGT	**Target**	**ADR Probability**			**Literature evidence**
		0-3 μM	**3-30 μM**	**>30 μM**	human (h), animal (a), *in vitro* (v)
cardiac arrhythmias	hERG (Binding)	-	0.03	0.002	h [26] a [27] v [28]
cardiac valve disorders	PDE3	0.05	-	0	h [30,31]
heart failures	hERG (Binding)	-	0.005	0	h [57]
myocardial disorders	GR (Binding)	0.02	-	0.005	h [32,33]
pericardial disorders	AdT	-	0.01	0	--

Table 2 Predicted renal ADR - target associations (detailed legend in Table 1).

Renal Disorder HLGT	**Target**	**ADR Probability**			**Literature evidence**
		0-3 μM	**3-30 μM**	**>30 μM**	human (h); animal (a), *in vitro* (v)
nephropathies	COX-2	0.003	-	0	h [34] a [35,36]
renal and urinary tract disorders congenital	PDE3	0.004	-	0	h [58] a [37,59]
renal disorders excl nephropathies	hERG (Binding)	-	0.01	0.0007	h [38] a [60]

Table 3 Predicted ADR associations with inhibition of the Bile Salt Export Pump (BSEP) transporter (detailed legend in Table 1).

HLGT	Target	ADR Probability			Literature evidence
		0-3 μM	**3-30 μM**	**>30 μM**	human (h); animal (a)
central nervous system vascular disorders	BSEP	-	0.09	0.008	(for BSEP and bile acid) a [61]
foetal complications	BSEP	0.01	-	0	h [45]
pregnancy labour delivery and postpartum conditions	BSEP	-	0.1	0	h [46]
lipid metabolism disorders	BSEP	-	0.2	0	h [62,63]
thyroid gland disorders	BSEP	-	0.07	0	a [64,65]
upper respiratory tract disorders excl infections	BSEP	0.1	-	0	h [66] a [67]
urolithiases	BSEP	-	0.07	0	h [68]
		0-30 μM	**30-300 μM**	**>300 μM**	
hepatic and hepatobiliary disorders	BSEP	-	0.2	0.09	h [43]

Acknowledgements

We are grateful to Mirjam Trame and Andy Stein for giving us the opportunity to participate in the 2018 Novartis Quantitative Sciences Academia-to-Industry Hackathon organized at the Novartis Institutes for Biomedical Research. We also thank Changchang Liu and Xinrui (Sandy) Zou for their contributions to the project at the Hackathon.

Author contributions

S.A., R.I., A.X.C., S.F., B.K., D.A. and L.U. conceived the study. D.A. and L.U. provided the Novartis *in vitro* pharmacology data, advice and mentorship. S.A., R.I., A.X.C., S.F. and B.K performed data analysis. S.A. developed the random forest modeling. R.I. developed the formalism for target-ADR association inference. S.F., R.I. and A.X.C. developed the query of OpenFDA. S.A., R.I., and A.X.C. wrote the paper and designed the figures with input from all the authors.

Conflict of interest

Authors declare no conflict of interest.

References

1. Institute of Medicine & Committee on Quality of Health Care in America. *To Err Is Human: Building a Safer Health System*. (National Academies Press, 2000).
2. Lazarou, J., Pomeranz, B. H. & Corey, P. N. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. *JAMA* **279**, 1200–1205 (1998).
3. Weiss, A. J., Freeman, W. J., Heslin, K. C. & Barrett, M. L. Adverse Drug Events in US Hospitals, 2010 Versus 2014. *HCUP Statistical Brief* **234**, (2018).
4. Lounkine, E. *et al.* Large-scale prediction and testing of drug activity on side-effect targets. *Nature* **486**, 361–367 (2012).
5. Bowes, J. *et al.* Reducing safety-related drug attrition: the use of in vitro pharmacological profiling. *Nat. Rev. Drug Discov.* **11**, 909–922 (2012).
6. Witek, R. P. & Bonzo, J. A. Perspective on In Vitro Liver Toxicity Models. *Applied In Vitro Toxicology* **4**, 229–231 (2018).
7. Whitebread, S., Hamon, J., Bojanic, D. & Urban, L. Keynote review: in vitro safety pharmacology profiling: an essential tool for successful drug development. *Drug Discov. Today* **10**, 1421–1433 (2005).
8. Liu, K. *et al.* Chemi-net: a graph convolutional network for accurate drug property prediction. *arXiv [cs.LG]* (2018).
9. Ekins, S. Predicting undesirable drug interactions with promiscuous proteins in silico. *Drug Discov. Today* **9**, 276–285 (2004).
10. Lynch, J. J., 3rd, Van Vleet, T. R., Mittelstadt, S. W. & Blomme, E. A. G. Potential functional and pathological side effects related to off-target pharmacological activity. *J. Pharmacol. Toxicol. Methods* **87**, 108–126 (2017).

11. Hamon, J. *et al.* In vitro safety pharmacology profiling: what else beyond hERG? *Future Med. Chem.* **1**, 645–665 (2009).
12. Mirams, G. R. *et al.* Simulation of multiple ion channel block provides improved early prediction of compounds' clinical torsadogenic risk. *Cardiovasc. Res.* **91**, 53–61 (2011).
13. Chen, A. W. Predicting adverse drug reaction outcomes with machine learning. *International Journal Of Community Medicine And Public Health* **5**, 901–904 (2018).
14. Nguyen, P. A., Born, D. A., Deaton, A. M., Nioi, P. & Ward, L. D. Phenotypes associated with genes encoding drug targets are predictive of clinical trial side effects. *bioRxiv* 285858 (2018). doi:10.1101/285858
15. U.S. Food and Drug Administration (FDA). Questions and Answers on FDA's Adverse Event Reporting System (FAERS). Available at: http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Surveillance/AdverseDrugEffects/default.htm.
16. Takarabe, M., Kotera, M., Nishimura, Y., Goto, S. & Yamanishi, Y. Drug target prediction using adverse event report systems: a pharmacogenomic approach. *Bioinformatics* **28**, i611–i618 (2012).
17. Mozzicato, P. MedDRA. *Pharmaceut. Med.* **23**, 65–75 (2009).
18. for Drug Statistics Methodology, W. C. C. Guidelines for ATC classification and DDD assignment. (2005).
19. U.S. Food and Drug Administration (FDA). openFDA. Available at: https://open.fda.gov.
20. Maciejewski, M. *et al.* Reverse translation of adverse event reports paves the way for de-risking preclinical off-targets. *Elife* **6**, (2017).
21. Sequeira, A. *et al.* Alpha 2A adrenergic receptor gene and suicide. *Psychiatry Res.* **125**, 87–93 (2004).
22. Cottingham, C. & Wang, Q. α2 adrenergic receptor dysregulation in depressive disorders: Implications for the neurobiology of depression and antidepressant therapy. *Neurosci.*

Biobehav. Rev. **36**, 2214–2225 (2012).

23. Drugs database, glyburide. Available at: https://www.drugs.com/glyburide.html.
24. Cerner Drug Database. Available at: https://www.cerner.com/solutions/drug-database.
25. Hanania, N. A. *et al.* The safety and efficacy of arformoterol and formoterol in COPD. *COPD* **7**, 17–31 (2010).
26. Curran, M. E. *et al.* A molecular basis for cardiac arrhythmia: HERG mutations cause long QT syndrome. *Cell* **80**, 795–803 (1995).
27. Wen, D., Liu, A., Chen, F., Yang, J. & Dai, R. Validation of visualized transgenic zebrafish as a high throughput model to assay bradycardia related cardio toxicity risk candidates. *J. Appl. Toxicol.* **32**, 834–842 (2012).
28. Mitcheson, J. S. hERG potassium channels and the structural basis of drug-induced arrhythmias. *Chem. Res. Toxicol.* **21**, 1005–1010 (2008).
29. Yusof, I., Shah, F., Hashimoto, T., Segall, M. D. & Greene, N. Finding the rules for successful drug optimisation. *Drug Discov. Today* **19**, 680–687 (2014).
30. Movsesian, M., Wever-Pinzon, O. & Vandeput, F. PDE3 inhibition in dilated cardiomyopathy. *Curr. Opin. Pharmacol.* **11**, 707–713 (2011).
31. Knight, W. & Yan, C. Therapeutic potential of PDE modulation in treating heart disease. *Future Med. Chem.* **5**, 1607–1620 (2013).
32. Aviña-Zubieta, J. A. *et al.* Immediate and past cumulative effects of oral glucocorticoids on the risk of acute myocardial infarction in rheumatoid arthritis: a population-based study. *Rheumatology* **52**, 68–75 (2013).
33. Oakley, R. H. & Cidlowski, J. A. Glucocorticoid signaling in the heart: A cardiomyocyte perspective. *J. Steroid Biochem. Mol. Biol.* **153**, 27–34 (2015).
34. Huerta, C., Castellsague, J., Varas-Lorenzo, C. & García Rodríguez, L. A. Nonsteroidal anti-inflammatory drugs and risk of ARF in the general population. *Am. J. Kidney Dis.* **45**, 531–539 (2005).

35. Wang, L. *et al.* Podocyte-specific knockout of cyclooxygenase 2 exacerbates diabetic kidney disease. *Am. J. Physiol. Renal Physiol.* **313**, F430–F439 (2017).
36. Slattery, P., Frölich, S., Schreiber, Y. & Nüsing, R. M. COX-2 gene dosage-dependent defects in kidney development. *Am. J. Physiol. Renal Physiol.* **310**, F1113–22 (2016).
37. Ye, H. *et al.* Modulation of Polycystic Kidney Disease Severity by Phosphodiesterase 1 and 3 Subfamilies. *J. Am. Soc. Nephrol.* **27**, 1312–1320 (2016).
38. Wadhwa, S., Wadhwa, P., Dinda, A. K. & Gupta, N. P. Differential expression of potassium ion channels in human renal cell carcinoma. *Int. Urol. Nephrol.* **41**, 251–257 (2009).
39. Fagerberg, L. *et al.* Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol. Cell. Proteomics* **13**, 397–406 (2014).
40. Nussey, S. S. & Whitehead, S. A. *Endocrinology: an integrated approach*. (CRC Press, 2013).
41. Miyamoto, J. *et al.* The pituitary function of androgen receptor constitutes a glucocorticoid production circuit. *Mol. Cell. Biol.* **27**, 4807–4814 (2007).
42. Nuzzi, R., Scalabrin, S., Becco, A. & Panzica, G. Gonadal Hormones and Retinal Disorders: A Review. *Front. Endocrinol.* **9**, 66 (2018).
43. Kubitz, R., Dröge, C., Stindt, J., Weissenberger, K. & Häussinger, D. The bile salt export pump (BSEP) in health and disease. *Clin. Res. Hepatol. Gastroenterol.* **36**, 536–553 (2012).
44. Riede, J., Poller, B., Huwyler, J. & Camenisch, G. Assessing the Risk of Drug-Induced Cholestasis Using Unbound Intrahepatic Concentrations. *Drug Metab. Dispos.* **45**, 523–531 (2017).
45. Liu, L.-Y., Wang, X.-H., Lu, Y., Zhu, Q.-R. & Wang, J.-S. Association of variants of ABCB11 with transient neonatal cholestasis : ABCB11 and TNC. *Pediatr. Int.* **55**, 138–144 (2013).
46. Geenes, V. & Williamson, C. Intrahepatic cholestasis of pregnancy. *World J. Gastroenterol.*

15, 2049 (2009).

47. Liu, M. *et al.* Large-scale prediction of adverse drug reactions using chemical, biological, and phenotypic properties of drugs. *J. Am. Med. Inform. Assoc.* **19**, e28–35 (2012).
48. Bender, A. *et al.* Analysis of pharmacology data and the prediction of adverse drug reactions and off-target effects from chemical structure. *ChemMedChem: Chemistry Enabling Drug Discovery* **2**, 861–873 (2007).
49. Dawson, S., Stahl, S., Paul, N., Barber, J. & Kenna, J. G. In vitro inhibition of the bile salt export pump correlates with risk of cholestatic drug-induced liver injury in humans. *Drug Metab. Dispos.* **40**, 130–138 (2012).
50. Montanari, F. *et al.* Flagging Drugs That Inhibit the Bile Salt Export Pump. *Mol. Pharm.* **13**, 163–171 (2016).
51. Tatonetti, N. P., Ye, P. P., Daneshjou, R. & Altman, R. B. Data-driven prediction of drug effects and interactions. *Sci. Transl. Med.* **4**, 125ra31 (2012).
52. Zhang, M. & Zhou, Z. A Review on Multi-Label Learning Algorithms. *IEEE Trans. Knowl. Data Eng.* **26**, 1819–1837 (2014).
53. Breiman, L. Random Forests. *Mach. Learn.* **45**, 5–32 (2001).
54. Charte, F. & Charte, D. Working with Multilabel Datasets in R: The mldr Package. *R J.* **7**, (2015).
55. Rivolli, A. & de Carvalho, A. C. The utiml Package: Multi-label Classification in R. *R J.* (2018).
56. Menze, B. H. *et al.* A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC Bioinformatics* **10**, 213 (2009).
57. Brugada, R. *et al.* Sudden death associated with short-QT syndrome linked to mutations in HERG. *Circulation* **109**, 30–35 (2004).
58. Pinto, C. S. *et al.* Phosphodiesterase Isoform Regulation of Cell Proliferation and Fluid

Secretion in Autosomal Dominant Polycystic Kidney Disease. *J. Am. Soc. Nephrol.* **27**, 1124–1134 (2016).

59. Wang, X., Ward, C. J., Harris, P. C. & Torres, V. E. Cyclic nucleotide signaling in polycystic kidney disease. *Kidney Int.* **77**, 129–140 (2010).
60. Babcock, J. J. & Li, M. hERG channel function: beyond long QT. *Acta Pharmacol. Sin.* **34**, 329–335 (2013).
61. Mertens, K. L., Kalsbeek, A., Soeters, M. R. & Eggink, H. M. Bile Acid Signaling Pathways from the Enterohepatic Circulation to the Central Nervous System. *Front. Neurosci.* **11**, 617 (2017).
62. Srivastava, A. Progressive familial intrahepatic cholestasis. *J. Clin. Exp. Hepatol.* **4**, 25–36 (2014).
63. Strautnieks, S. S. *et al.* A gene encoding a liver-specific ABC transporter is mutated in progressive familial intrahepatic cholestasis. *Nat. Genet.* **20**, 233–238 (1998).
64. Watanabe, M. *et al.* Bile acids induce energy expenditure by promoting intracellular thyroid hormone activation. *Nature* **439**, 484–489 (2006).
65. Mukaisho, K.-I. *et al.* High serum bile acids cause hyperthyroidism and goiter. *Dig. Dis. Sci.* **53**, 1411–1416 (2008).
66. Comhair, S. A. A. *et al.* Metabolomic Endotype of Asthma. *J. Immunol.* **195**, 643–650 (2015).
67. Tan, X. *et al.* Genetic and Proteomic characterization of Bile Salt Export Pump (BSEP) in Snake Liver. *Sci. Rep.* **7**, 43556 (2017).
68. Li, C.-H., Sung, F.-C., Wang, Y.-C., Lin, D. & Kao, C.-H. Gallstones increase the risk of developing renal stones: a nationwide population-based retrospective cohort study. *QJM* **107**, 451–457 (2014).