Received: 26 July 2018 | Revised: 10 September 2018 | Accepted: 21 September 2018

DOI: 10.1002/jat.3741

REVIEW ARTICLE

WILEY Journal of Applied Toxicology

Prediction of clinically relevant drug-induced liver injury from structure using machine learning

Felix Hammann | Verena Schöning | Jürgen Drewe

Division of Clinical Pharmacology and Toxicology, University Hospital Basel, Basel, Switzerland

Correspondence
Felix Hammann, Divison of Clinical Pharmacology and Toxicology, University Hospital of Basel, Schanzenstr. 55, CH 4031 Basel, Switzerland.
Email: felix.hammann@unibas.ch

Abstract

Drug-induced liver injury (DILI) is the most common cause of acute liver failure and often responsible for drug withdrawals from the market. Clinical manifestations vary, and toxicity may or may not appear dose-dependent. We present several machine-learning models (decision tree induction, k-nearest neighbor, support vector machines, artificial neural networks) for the prediction of clinically relevant DILI based solely on drug structure, with data taken from published DILI cases. Our models achieved corrected classification rates of up to 89%. We also studied the association of a drug's interaction with carriers, enzymes and transporters, and the relationship of defined daily doses with hepatotoxicity. The results presented here are useful as a screening tool both in a clinical setting in the assessment of DILI as well as in the early stages of drug development to rule out potentially hepatotoxic candidates.

KEYWORDS

drug-induced liver injury, hepatotoxicity, machine learning, network analysis, structure-activity relationships

1 | INTRODUCTION

Drug-induced liver injury (DILI) is a diagnosis of exclusion for hepatotoxicity causally linked to a xenobiotic (synthetic drugs, herbal preparations, dietary supplements) when all other explanations have been ruled out. It is the most common cause of acute liver failure in developed countries, and a major reason for withdrawal of approved drugs from the US market (Lasser et al., 2002; Reuben, Koch, Lee, & Acute Liver Failure Study, 2010). The manifestations range from asymptomatic elevation of liver enzymes to outright acute liver failure. The two main clinical pictures are hepatocellular damage and cholestasis, with many intermediate presentations, as well as changes as liver damage progresses and resolves (Benichou, Danan, & Flahault, 1993; Danan & Benichou, 1993). This heterogeneity is reflected by the various forms of pathophysiological mechanisms implicated, which include disruption of mitochondrial metabolism, changes in transport protein function, immunological processes and hypersensitivity, and direct hepatocellular damage (Kock et al., 2014). Antibiotics are a common source of DILI, with amoxicillin/clavulanic acid posing the greatest risk.

Risk factors are bioactivation by metabolic enzymes (Boelsterli & Lee, 2014; Thompson, Isin, Ogese, Mettetal, & Williams, 2016), higher lipophilicity (logP $\geq$3) and dose (daily dose $\geq$50 mg) (Chalhoub, Sliman, Arumuganathan, & Lewis, 2014; Chen, Borlak, & Tong, 2013; Yu et al., 2014). In addition, DILI has been observed after low-dose medications (Lammert et al., 2008) in patients with a predisposition due to genetic polymorphisms or other ADMET particularities that have gone unrecognized until now, resulting in a false labeling of an adverse event as idiosyncratic.

A well-described risk factor for causing cholestatic injury is the inhibition of the canalicular bile salt export pump. Hepatocytes are thought to be flooded with bile salts, eventually leading to apoptosis (Morgan et al., 2010). The basolateral ATP-dependent efflux pumps MRP3 (ABCC3) and MRP4 (ABCC4) can be recruited to shift bile salts into the sinusoidal veins (they are in fact upregulated in cholestasis), and inhibition can among other factors (Aleo et al., 2014; Guo et al., 2015) contribute to cholestatic DILI (Chai et al., 2012; Gradhand et al., 2008). Immunological processes have also been shown to play a role in flucloxacillin cholestatic DILI, wherein hepatic biliary cells are destroyed preferentially in HLA-B*5701-positive patients (Daly et al., 2009).

Previous research, showing a higher risk for certain DILIs in specific countries or ethnicities, supports the existence of a genetic

component (Ibanez, Perez, Vidal, Laporte,, & Grup d'Estudi Multicenteric d'Hepatotoxicitat Aguda de, 2002). In addition, females appear to be more susceptible to DILI than males (Parkinson, Mudra, Johnson, Dwyer, & Carroll, 2004).

While risk factors can predispose an individual to develop DILI, these risk factors are often not known, and such cases are then often labeled as "idiosyncratic." However, to assess DILI clinically, drug-related risk factors also need to be taken into account,e.g., certain structural motifs or other physicochemical properties. To the best of our knowledge, there is still a lack of a predictive model for clinically manifest DILI, a tool that could be a valuable adjunct in evaluating hepatic dysfunction in a given patient.

A drug's defined daily dose (DDD) is a standardized measure of drug consumption. Interestingly, it appears that high daily doses are predictive of DILI, particularly when administered with cytochrome P450 inhibitors (Chen et al., 2013; Yu et al., 2014). The respective authors believed this to be the result of an increased exposure to mother substances of a drug through both higher dose and decreased detoxification. Another possibility is that more complex (i.e., heavier) drugs have greater hepatotoxic potential. An analysis of molar DDDs could answer this question.

2 | METHODS

2.1 | Data collection and preparation

2.1.1 | Data acquisition and structure analysis

The datasets of DILI-positive compounds were taken from different sources, consisting of 311 drugs, which were withdrawn from the market in the USA (Ekins, Williams, & Xu, 2010) or European countries due to hepatotoxicity, not marketed there, have received a black box warning because of hepatotoxicity, or are well-known hepatotoxic agents. Other sources were literature-based databases (Ekins et al., 2010; Greene et al., 2010; Stine & Lewis, 2011), and 319 drugs from the three Western DILI registries (USA, Sweden and Spain) (Stine & Lewis, 2011). We found a total of 627 individual substances in the literature. From these, we removed ambiguous identifiers (e.g., "estrogens").

We also removed proteins and peptides as well as metallic or inorganic compounds (e.g., arsenic trioxide, iron sulfate). This restricted our dataset to one of small molecule substances chemically similar to what is used in most areas of pharmacotherapy today. Furthermore, the structural and physicochemical parameters calculated in this study are largely applicable only to smaller molecules with a unique structure. We used the PubChem Substance and Compounds databases (http://pubchem.ncbi.nlm.nih.gov/) to find the associated two-dimensional structures in a simplified molecular-input line-entry system (SMILES; isomeric if available, canonical otherwise). Finally, we stripped the molecules of associated salts under the assumption that they are pharmacologically inert. This process ultimately gave us a list of 588 compounds labeled either "hepatotoxic" or "non-hepatotoxic."

Initial physicochemical calculations were performed with the PaDEL-Descriptor package (version 2.21). We computed the entire range of available 1D and 2D descriptors (n = 1381) for all compounds. As some descriptors cannot be calculated for all molecules for technical reasons, this resulted in 526 complete cases (i.e., molecules with complete sets of descriptors). Most incomplete cases were due to only a select few descriptors. We therefore excluded all descriptors that failed in 5% of molecules, which brought the number of complete cases to n = 575. The descriptors removed (n = 63), included several eigenvalues of the Burden matrix (BCUT) (Burden, 1989), simple and valence chi chain descriptors (SCH, VCH), valence and average valence path descriptors (VP, AVP), and a van-der-Waals volume descriptor (VABC).

The remaining incomplete cases were gallium nitrate, trichloroethylene, bromoethanamine, sodium bicarbonate, carbon tetrachloride, chloroform, cadmium chloride, thioacetamide, probucol, dichloroethylene, hydrazine, nitrosamine and ferrous sulfate. We removed them, as they are not representative of small molecule drugs. Afterwards, we removed low-variance descriptors, which were mostly counts of substructural motifs. The final set consisted of 575 compounds and 1001 descriptors.

For metabolic information, we turned to DrugBank Version 4.3 (https://www.drugbank.ca). DrugBank is a freely available resource maintained by the University of Alberta, Canada, which, among other things, provides curated information on drug targets and metabolic pathways. We downloaded the entire database and constructed the network of drugs to bioentities (BE; an umbrella term comprising metabolic enzymes, transporters, carriers and targets). From this network we removed all substances not in our dataset as well as BE that had no association with the remaining substances (i.e., if an enzyme did not interact with any of the compounds in our dataset, it was deleted from the network). A total of 417 substances (70.9%) were listed in DrugBank. Because some of the interactions are asymmetrical (drug to target) and some are not (a drug can be metabolized by and/or induce/inhibit an enzyme's activity) we chose an undirected network architecture. The network is also bipartite as no drugs were assumed to interact directly with each other, and the same assumption was made for BEs. We then constructed unipartite projections so that drugs are removed from the network, and edges (connections) were inserted where two BEs interact with the same drug. For example, the lipid lowering drug simvastatin is a substrate of both cytochrome P450 CYP3A4 and CYP2D6. This would correspond to a connection (edge) between the two isoforms when simvastatin is removed. We performed these steps separately for hepatotoxic and non-hepatotoxic compounds, leaving us with two different networks that can help understand differences in metabolism in DILI and non-DILI situations. The complete architecture is given in the supporting information.

2.1.2 | Structural similarity

The structural heterogeneity of a collection of molecules can be quantified by considering individual molecules as points in a high-dimensional space wherein each axis corresponds to a descriptor. Similar compounds will then lie closer together, and a set of compounds is considered homogeneous if it is tightly packed. The Tanimoto coefficient is a widely adopted method, where the similarity between compounds i and j is calculated from a set of k descriptors as.

$$sim(i,j) = \frac{\sum_{d=1}^{k} X_{di}X_{dj}}{\sum_{d=1}^{k} (X_{di})^2 + \sum_{d=1}^{k} (X_{dj})^2 - \sum_{d=1}^{k} X_{di}X_{dj}}$$

The values of the coefficient range from 0 to 1, with low values indicating diversity and high values similarity.

2.1.3 | Model learning process

All models discussed here were learned with 10-fold cross-validation to avoid overfitting. Overfitting arises when models with high degrees of complexity and a high accuracy are created that are not generalizable,i.e., perform much worse on unseen data. Additionally, we repeated the cross-validated learning runs 10 times with different random seeds to detect any variations in model quality. The final reported models were chosen from these 10 runs.

We judged model performance based on their corrected classification rate (CCR), given as.

$$CCR = \frac{1}{2}\left(\frac{T_N}{N_0} + \frac{T_P}{N_1}\right)$$

for the two-class case. T_N and T_P represent the number of true negative and positive predictions, respectively, and N_0 and N_1 the total number of negative and positive observations in the model. This measure is more appropriate for skewed datasets such as the one presented here where one class (hepatotoxic compounds) outnumbers the other (non-hepatotoxic compounds).

We surveyed several commonly used machine-learning paradigms: decision tree induction (DTI), k-nearest neighbor classification (kNN), support vector machines (SVM) and artificial neural networks (ANN). We implemented these models in GNU R Version 3.3.3.

DTI is not considered to require feature selection as the number of attributes included in the models is limited by the learning parameters (e.g., maximal tree depth, minimum number of instances per split, minimum number of instances per node). For other paradigms (kNN, SVM, ANN), we performed separate feature selection (dimensionality reduction) with two commonly used methods: recursive feature elimination (RFE) and correlation-based feature subset selection (CFSS). We provide a full list of the descriptors selected by each method in the supporting information.

As a last step, we repeated the model building processes with *y*-randomization (Rücker, Rücker, & Meringer, 2007). Here, the observed activities were replaced with random activities with the same proportions of classes as the original data. This is useful to ensure models detect true relationships between attributes and outcomes in situations where the number of attributes and the dimensionality of the paradigms (which can equal infinity in SVM setups) are very large.

2.1.4 | Defined daily doses

The WHO Collaborating Centre for Drug Statistics Methodology maintains a list of drugs and their DDDs (https://www.whocc.no/atc_ddd_index). We manually checked the 588 substances in our original dataset against this database and noted the maximum DDD. No DDD was recorded when the mode of application was topical or local (creams, inhalers, etc.), assuming that no systemic exposure (and, consequently, hepatotoxicity) occurs with their use. We found 245 (41.6%) drugs for which we recorded the dose in mg/day and the millimolar dose (mmol/day; conversion made with molecular mass as per PaDEL calculations).

2.1.5 | Software

Gnu R 3.3.3 (http://www.R-project.org), Gephi 0.82 (https://gephi.org), PaDEL-Descriptor 2.21 (http://www.yapcwsoft.com/dd/padeldescriptor) and IBM SPSS Statistics version 25 were used in this study.

3 | RESULTS

3.1 | Dataset

The final set for the creation of the machine-learning models contains 384 (66.8%) DILI-positive drugs and 191 (33.2%) DILI-negative drugs (total n = 575), and is reproduced in the Supporting information. The overall Tanimoto similarity index value was fairly low at 0.24, indicating a heterogeneous dataset based on the descriptors employed.

3.2 | Decision tree induction

We performed decision tree analysis with implementation of the CART algorithm in GNU R ("r-part"). The minimum number of cases per split was set to 10, and the minimum number of instances per node was set to 5. Models were learned from the original data with 10-fold cross-validation. The final model performed with a CCR of 0.89 and is reproduced in Figure 1. *Y*-randomized runs had a maximum CCR of 0.53. There was no increase in performance by balancing datasets during the learning process (maximum CCR 0.88).

The descriptors selected in this model were mostly topological and include autocorrelation descriptors (AATS2e, AATS2m, AATS4p, AATS5m, AATSC1c, ATS2e, ATSC0e, ATSC3e, ATSC3v, ATSC4e, ATSC4s, ATSC4v, ATSC6v, MATS1e, MATS3c), atom type electrotopological state descriptors (hmin, maxaasC, maxsNH2, SHBa) (Gramatica, Corradi, & Consonni, 2000; Hall & Kier, 1995), structural information content (an index of neighborhood symmetry of the third order, SIC3) (Basak, Harriss, & Magnuson, 1984), the topological distance matrix (SpMin1_Bhs, SpMax1_Bhi), Barysz matrix (VE1_D, VE1_Dzs) (Barysz, Jashari, Lall, Srivastava, & Trinajstic, 1983) and molecular polarizability (Mp). All of these descriptors serve to characterize different molecular shapes, branching and distributions of charge.

The most readily interpretable attributes were an estimator of logP (ALogP) with a cut-off of −0.72, where higher values are more likely to be predicted as hepatotoxic, and the number of hydrogens (nH). The latter appears very late (i.e., the decision influences few compounds), with >20 hydrogens being associated with hepatotoxicity.

3.3 | k-Nearest neighbors

We screened several values of k (5-20) and found the best performance for k = 11. The CCR was 0.73, although little difference was

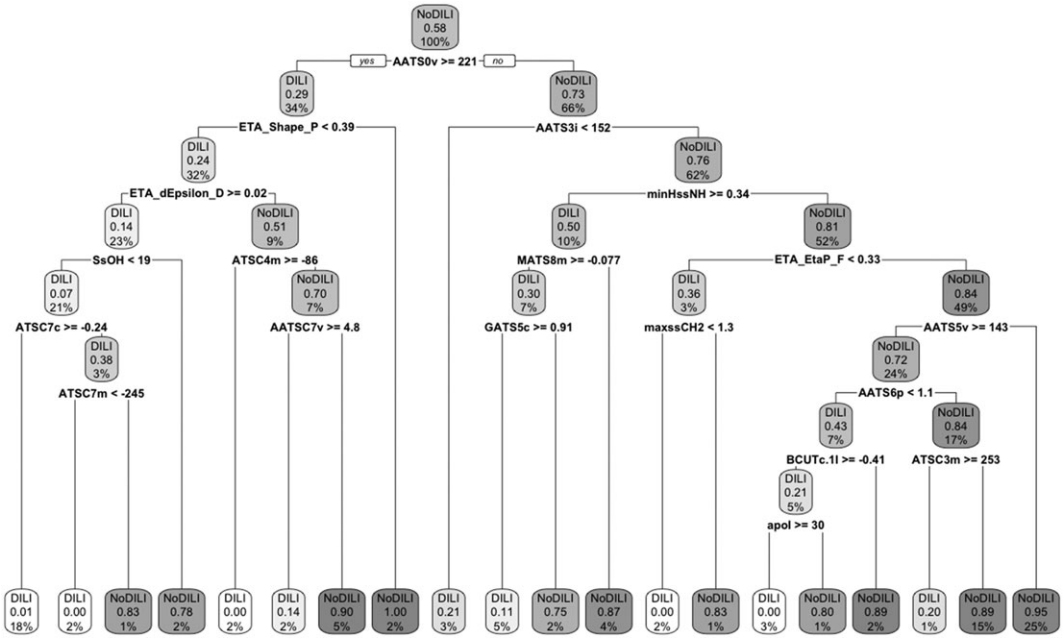

FIGURE 1 Decision tree model for hepatotoxic ("DILI") and non-hepatotoxic ("NoDILI") compounds. More intense the shading, the more of one class is present at each node. DILI, drug-induced liver injury

seen between different *k* values (minimal CCR = 0.71). The descriptor set used here (*n* = 27) was the one selected in the DTI model. Other feature subset selection methods (RFE, CFSS) were markedly less successful (maximum CCR = 0.65).

We were able to increase the predictive performance on the original dataset to a CCR of 0.83 (maximum CCR in *y*-randomized runs was 0.56) by using SMOTE balanced internal training sets during cross-validation. Again, *k* = 11 produced the best model.

3.4 | Support vector machines

The CCR of the best performing SVM model was 0.74 (CCR = 0.54 in *y*-randomization) for the decision tree feature subset, while RFE and CFSS subsets were less successful (maximum CCR = 0.66). Using balanced datasets markedly increased the CCR to 0.98, with specificity and sensitivity at equally high values (0.98). The CCR in the *y*-randomized runs was 0.89, however, which is why we chose to discard the models learned using balanced training data.

3.5 | Artificial neural networks

We trained feed-forward neural networks with a single hidden layer with all feature sets. The best performance was seen with the decision tree feature set (CCR = 0.86, CCR in *y*-randomization = 0.49), while RFE and CFSS both achieved CCRs of 0.74. Balancing the training data did not improve predictivity.

3.6 | Interactions with bioentities

Our survey of DrugBank listed interactions with carriers, transporters and metabolizing enzymes showed (Figure 2) that the largest share of interactions were with CYP3A4, CYP2C9, MRP1, CYP2D6 and CYP2C19. Of statistical significance were CYP2C9, CYP2C8, CYP3A5/7, SLC22A6, ABCC2, serum albumin and prostaglandin G/H synthase 1. It is of particular interest, that there is not only a statistical difference between individual BE but that the network of interactions (see Supporting information) is more complex for hepatotoxic compounds compared to non-toxic compounds.

3.7 | Defined daily doses

Based on previous analysis (Chalhoub et al., 2014; Chen et al., 2013; Yu et al., 2014), the distribution of compounds' DDDs with regard to the threshold of 50 mg between hepatotoxic and non-hepatotoxic groups showed (Figure 3) that for compounds with DDD <50 mg the same number of compounds have been observed to be hepatotoxic and non-hepatotoxic (*n* = 48 each). For compounds with DDD ≥50 mg, more than twice as many compounds were hepatotoxic (*n* = 167) than non-hepatotoxic (*n* = 68; *P* < 0.001). An ROC analysis showed that the criterion of DDD ≥50 mg alone was, however, only a moderate predictor of hepatotoxicity with a sensitivity of 77.2% and specificity of 22.4%. The same sensitivity was observed for micromolar DDD (threshold 144 μmol) with a specificity of 40.5% and a significant (*P* = 0.001) distribution of DDD between the treatment groups.

4 | DISCUSSION

4.1 | Predictive models

Our survey of machine-learning paradigms for the prediction of clinically relevant hepatotoxicity points to decision trees as the most useful method. Its precision is matched by ANNs and, if combined with the more involved cross-validation process of balancing the training data, also by the kNN method. DTI also provided the most informative feature selection for paradigms that require dimensionality reduction (such as SVM, ANN or kNN). Because

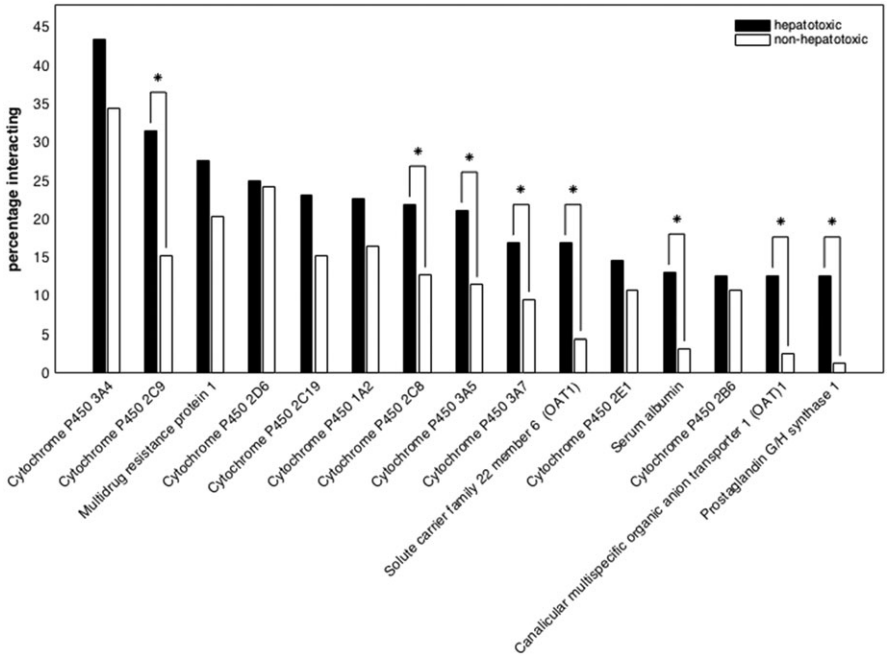

FIGURE 2 Fraction of drugs interacting with the 15 most common enzymes, carriers, transporters and targets, grouped by hepatotoxicity. Significance brackets (*) at $P < 0.05$ (Fisher's exact test)

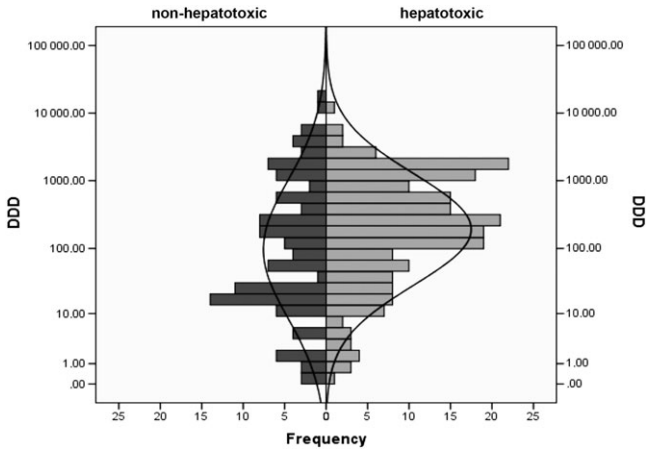

FIGURE 3 Distribution of DDDs is different for hepatotoxic and non-hepatotoxic compounds ($P < 0.001$). DDD, defined daily dose

hepatotoxicity can occur for a variety of reasons, the capability of DTI to separate a problem space (DILI) into several subspaces (different pathomechanisms) could be grounds for its effectiveness in this setting.

To put our results into context, we evaluated comparable machine-learning approaches to modeling DILI in literature published from 2011 to 2017 (full reference list and overview given as supporting information). Of the 15 studies, some were performed solely on animal data. Of those based on human data, only five used clinically validated outcomes (such as the FDA's Liver Toxicity Knowledge Base) that can be held to a similarly rigorous standard. Our DTI models are highly predictive, and are on par with—if not superior to—the other published efforts.

Not only can hepatotoxicity arise through a multitude of mechanisms but it can also be precipitated by risk factors. For example, age and gender can greatly affect liver function and toxicity of compounds by changes in cytochrome activity, reduction in hepatic blood flow, decreased drug binding and malnutrition. (Hunt, Westerkam, & Stave, 1992) Valproate and erythromycin, for instance, show greater hepatotoxicity in children compared to adults. Ethanol consumption increases toxicity of acetaminophen through induction of CYP2E1 (chronic intake) and formation of *N*-acetyl-p-benzoquinone imine as well as reduction of glutathione stores necessary for *N*-acetyl-p-benzoquinone imine elimination (Stine & Chalasani, 2017). Isoniazid toxicity through slow acetylation is genetically predetermined and racial predisposition has been extensively researched (Walker et al., 2009). Risk also increases with concomitant diseases (e.g., pre-existing liver disease, diabetes mellitus, renal failure, HIV/AIDS, obesity) or drug-drug interactions. These individual host factors are not represented in structure-activity relationship models. This is of course complicated by the fact that the same drug can induce multiple forms of injury, in certain cases at different time points in therapy (early hepatocellular damage progressing to mixed patterns). Therefore, it is possible that even better predictivity could be achieved by integrating individual patient characteristics into future models,e.g., disease state, genotype or concomitant medications.

4.2 | Interactions with bioentities

From our analysis of DrugBank, we found that the network of interactions for hepatotoxic compounds appears more interconnected for hepatotoxic compounds. There were also several statistically significant interaction differences for both classes. It is important to note that DrugBank's data is the result of extensive literature searches and not systematic in vitro testing of the interactions given. As such, there are bound to be biases. For instance, routine regulatory and industry endpoints such as CYP3A4 interactions will be more consistently determined than other endpoints that may stem from academic research.

4.3 | CYP3A subfamily

Members of the CYP3A subfamily are involved in the phase I metabolism of an estimated 50% of small molecule drugs on the market. They catalyze a variety of reactions such as dealkylations, hydroxylations, aromatic oxidations, dehydrogenations and epoxidations (Rendic, 2002). The last process specifically produces exquisitely toxic compounds: the highly reactive electrophilic epoxides (Niederer, Behra, Harder, Schwarzenbach, & Escher, 2004). Like members of the CYP2C family, CYP3A4 has epoxygenase activity. This has been shown for arachidonic acid and its epoxygeneration to carcinogenous compounds (Bishop-Bailey, Thomson, Askari, Faulkner, & Wheeler-Jones, 2014). There are also specific examples of how CYP3A4 metabolism confers toxicity. The *N*-demethylation of cocaine is greater with CYP3A4 induction and so is cocaine's hepatotoxicity (Pellinen et al., 1994). While there seems to be a trend towards CYP3A4 interactions for hepatotoxic compounds, there were only statistical significances for CYP3A5 and CYP3A7. All of these enzymes have interindividual variability in activity and, in the case of CYP3A4 and CYP3A7, they show developmental patterns, with CYP3A4 increasing in activity through infancy and CYP3A7 losing activity with age (de Wildt, Kearns, Leeder, & van den Anker, 1999). Possibly, this variability explains why CYP3A4 did not reach significance. Another explanation could be the substrate overlap for CYP2C8 and CYP3A4.

4.4 | CYP2C subfamily

The isoforms 2C8, 2C9, 2C18 and 2C19 share >80% amino acid identity, although there is only little overlap in substrates. The subfamily accounts for about 20% of cytochrome activity in hepatic microsomes. Whereas CYP2C8 metabolizes weakly acidic large molecules, CYP2C9 recognizes weak acids with a hydrogen acceptor, and CYP2C19 basic molecules or amides with two hydrogen acceptors (Zanger, Turpeinen, Klein, & Schwab, 2008). Many have narrow therapeutic indices (coumarines, phenytoin). The CYPs 2C8, 2C9, 2C18 and 2C19 have epoxygenase activity and can generate superoxide (O_2^-), cytotoxic reactive oxygen species (Fleming, 2014; Miners & Birkett, 1998).

4.5 | CYP2D6

The CYP2D6 enzyme is polymorphic,i.e., there are several interindividual and ethnic differences in activity, and individuals can be grouped into different phenotypes (ultra-rapid, extensive, intermediate and poor metabolizers). There are inhibitors,e.g., fluoxetine and quinidine, but no known inducers of CYP2D6 (Teh & Bertilsson, 2012). This variation has drug safety consequences, as many antipsychotics and antidepressants, but also oncologicals,e.g., tamoxifene, are metabolized at least partly by CYP2D6. As a consequence, there may be symptoms of overdose where the dosage is not matched with individual metabolic capacity, or where there are toxic metabolites. This has been documented in vitro and in vivo for psychiatric drugs (e.g., quetiapine, venlafaxine and trazodone) (Jornil et al., 2013; Li & Cameron, 2012; Najibi et al., 2016). Other examples include primaquine (Ganesan, Tekwani, Sahu, Tripathi, & Walker, 2009) and acetaminophen (Dong, Haining, Thummel, Rettie, & Nelson, 2000). It is therefore not surprising that CYP2D6 interactions are a risk factor for drug toxicity.

4.6 | Serum albumin

Binding to albumin, alpha-1 acid glycoprotein and lipoproteins can play an important part in drug distribution. Hypoproteinemia will lead to higher free (= unbound) concentrations for these drugs, and, by consequence, stronger effects and possibly toxicity. Clinically, this is well recognized for antiepileptics such as valproic acid (Ahmed & Siddiqi, 2006), anti-inflammatory drugs such as salicylates (Gitlin, 1980) or anti-infectives (Makhlouf, Helmy, Fawzy, El-Attar, & Rashed, 2008). Many disease states can influence protein binding, ranging from malnutrition and malignancies to hepatotoxicity itself, as hepatic insufficiency involves altered protein synthesis.

4.7 | Cellular transporters

The basolaterally expressed MRP1 (ABCC1) and the apically expressed MRP2 (ABCC2) are members of the ATP-binding cassette family, found throughout the body (also hepatically), and involved in the transport of a wide range of compounds, both charged and uncharged. Their clinical importance can be seen in the response to and toxicity of methotrexate depending on MRP1 activity (Lima, Bernardes, Azevedo, Medeiros, & Seabra, 2015), or statin disposition in relation to MRP1 and MRP2 activity (Rodrigues, 2010).

4.8 | Defined daily doses

We confirmed previous observations (Chalhoub et al., 2014; Chen et al., 2013; Yu et al., 2014) that DDD is a predictor of hepatotoxicity. The trend towards better specificity of high DDDs in micromolar units compared to milligrams is indication that the sheer number of circulating molecules is more important than the dose amount in the system. However, DDDs alone were both only moderate predictors of hepatotoxicity with sensitivity 72.2% and low specificity.

5 | CONCLUSIONS

We present a study of the structural and metabolic features associated with hepatotoxicity. There are only few instances where DILI is not considered idiosyncratic. Our study indicates that this is not the case, and that the vast majority of hepatotoxicity seems to be predictable from a drug's structure—a potentially very useful tool in clinical pharmacological practice as well for avoiding costly attritions in drug development.

Despite the predictive power of our models, they could be markedly improved by incorporating these susceptibility factors into more comprehensive systems. These could help in individual therapy (personalized medicine) and in regulatory questions,e.g., for judging the toxic potential in special populations such as children, the elderly or pregnant/lactating women. Similarly, given a sufficiently large dataset, subgroup analyses by injury pattern (hepatocellular vs. cholestatic vs. mixed) would be informative.

We also show that different metabolic pathways are active in hepatotoxicity, and these may be influenced by predisposing factors (age, gender or ethnicity), concomitant medication or disease states. The major limitation here is that the drug interactions

evaluated are likely heavily biased by regulatory requirements. Furthermore, we were able to confirm that higher DDDs, particularly in micromolar units, are a risk factor for the development of acute hepatotoxic effects.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to report.

ORCID

Felix Hammann http://orcid.org/0000-0003-0658-9330

REFERENCES

Ahmed, S. N., & Siddiqi, Z. A. (2006). Antiepileptic drugs and liver disease. *Seizure*, *15*(3), 156–164. https://doi.org/10.1016/j.seizure.2005.12.009

Aleo, M. D., Luo, Y., Swiss, R., Bonin, P. D., Potter, D. M., & Will, Y. (2014). Human drug-induced liver injury severity is highly associated with dual inhibition of liver mitochondrial function and bile salt export pump. *Hepatology (Baltimore, Md)*, *60*(3), 1015–1022. https://doi.org/10.1002/hep.27206

Barysz, M., Jashari, G., Lall, R. S., Srivastava, A. K., & Trinajstic, N. (1983). On the distance matrix of molecules containing heteroatoms. In *Chemical Applications of Topology and Graph Theory* (pp. 222–230). Amsterdam, The Netherlands: Elsevier.

Basak, S. C., Harriss, D. K., & Magnuson, V. R. (1984). Comparative study of lipophilicity versus topological molecular descriptors in biological correlations. *Journal of Pharmaceutical Sciences*, *73*(4), 429–437. https://doi.org/10.1002/jps.2600730403

Benichou, C., Danan, G., & Flahault, A. (1993). Causality assessment of adverse reactions to drugs—II. An original model for validation of drug causality assessment methods: case reports with positive rechallenge. *Journal of Clinical Epidemiology*, *46*(11), 1331–1336. https://doi.org/10.1016/0895-4356(93)90102-7

Bishop-Bailey, D., Thomson, S., Askari, A., Faulkner, A., & Wheeler-Jones, C. (2014). Lipid-metabolizing CYPs in the regulation and dysregulation of metabolism. *Annual Review of Nutrition*, *34*, 261–279. https://doi.org/10.1146/annurev-nutr-071813-105747

Boelsterli, U. A., & Lee, K. K. (2014). Mechanisms of isoniazid-induced idiosyncratic liver injury: emerging role of mitochondrial stress. *Journal of Gastroenterology and Hepatology*, *29*(4), 678–687. https://doi.org/10.1111/jgh.12516

Burden, F. R. (1989). Molecular identification number for substructure searches. *Journal of Chemical Information and Computer Sciences*, *29*(3), 225–227. https://doi.org/10.1021/ci00063a011

Chai, J., He, Y., Cai, S. Y., Jiang, Z., Wang, H., Li, Q., … Chen, W. (2012). Elevated hepatic multidrug resistance-associated protein 3/ATP-binding cassette subfamily C 3 expression in human obstructive cholestasis is mediated through tumor necrosis factor alpha and c-Jun NH2-terminal kinase/stress-activated protein kinase-signaling pathway. *Hepatology*, *55*(5), 1485–1494. https://doi.org/10.1002/hep.24801

Chalhoub, W. M., Sliman, K. D., Arumuganathan, M., & Lewis, J. H. (2014). Drug-induced liver injury: what was new in 2013? *Expert Opinion on Drug Metabolism & Toxicology*, *10*(7), 959–980. https://doi.org/10.1517/17425255.2014.909408

Chen, M., Borlak, J., & Tong, W. (2013). High lipophilicity and high daily dose of oral medications are associated with significant risk for drug-induced liver injury. *Hepatology (Baltimore, Md)*, *58*(1), 388–396. https://doi.org/10.1002/hep.26208

Daly, A. K., Donaldson, P. T., Bhatnagar, P., Shen, Y., Pe'er, I., Floratos, A., … International SAE Consortium (2009). HLA-B*5701 genotype is a major determinant of drug-induced liver injury due to flucloxacillin. *Nature Genetics*, *41*(7), 816–819. https://doi.org/10.1038/ng.379

Danan, G., & Benichou, C. (1993). Causality assessment of adverse reactions to drugs—I. A novel method based on the conclusions of international consensus meetings: application to drug-induced liver injuries. *Journal of Clinical Epidemiology*, *46*(11), 1323–1330. https://doi.org/10.1016/0895-4356(93)90101-6

de Wildt, S. N., Kearns, G. L., Leeder, J. S., & van den Anker, J. N. (1999). Cytochrome P450 3A: ontogeny and drug disposition. *Clinical Pharmacokinetics*, *37*(6), 485–505. https://doi.org/10.2165/00003088-199937060-00004

Dong, H., Haining, R. L., Thummel, K. E., Rettie, A. E., & Nelson, S. D. (2000). Involvement of human cytochrome P450 2D6 in the bioactivation of acetaminophen. *Drug Metabolism and Disposition*, *28*(12), 1397–1400.

Ekins, S., Williams, A. J., & Xu, J. J. (2010). A predictive ligand-based Bayesian model for human drug-induced liver injury. *Drug Metabolism and Disposition*, *38*(12), 2302–2308. https://doi.org/10.1124/dmd.110.035113

Fleming, I. (2014). The pharmacology of the cytochrome P450 epoxygenase/soluble epoxide hydrolase axis in the vasculature and cardiovascular disease. *Pharmacological Reviews*, *66*(4), 1106–1140. https://doi.org/10.1124/pr.113.007781

Ganesan, S., Tekwani, B. L., Sahu, R., Tripathi, L. M., & Walker, L. A. (2009). Cytochrome P(450)-dependent toxic effects of primaquine on human erythrocytes. *Toxicology and Applied Pharmacology*, *241*(1), 14–22. https://doi.org/10.1016/j.taap.2009.07.012

Gitlin, N. (1980). Salicylate hepatotoxicity: the potential role of hypoalbuminemia. *Journal of Clinical Gastroenterology*, *2*(3), 281–285. https://doi.org/10.1097/00004836-198009000-00018

Gradhand, U., Lang, T., Schaeffeler, E., Glaeser, H., Tegude, H., Klein, K., … Fromm, M. F. (2008). Variability in human hepatic MRP4 expression: influence of cholestasis and genotype. *The Pharmacogenomics Journal*, *8*(1), 42–52. https://doi.org/10.1038/sj.tpj.6500451

Gramatica, P., Corradi, M., & Consonni, V. (2000). Modelling and prediction of soil sorption coefficients of non-ionic organic pesticides by molecular descriptors. *Chemosphere*, *41*(5), 763–777. https://doi.org/10.1016/S0045-6535(99)00463-4

Greene, N., Fisk, L., Naven, R. T., Note, R. R., Patel, M. L., & Pelletier, D. J. (2010). Developing structure-activity relationships for the prediction of hepatotoxicity. *Chemical Research in Toxicology*, *23*(7), 1215–1222. https://doi.org/10.1021/tx1000865

Guo, Y. X., Xu, X. F., Zhang, Q. Z., Li, C., Deng, Y., Jiang, P., … Peng, W. X. (2015). The inhibition of hepatic bile acids transporters Ntcp and Bsep is involved in the pathogenesis of isoniazid/rifampicin-induced hepatotoxicity. *Toxicology Mechanisms and Methods*, *25*(5), 382–387. https://doi.org/10.3109/15376516.2015.1033074

Hall, L. H., & Kier, L. B. (1995). Electrotopological state indices for atom types: A novel combination of electronic, topological, and valence state information. *Journal of Chemical Information and Computer Sciences*, *35*(6), 1039–1045. https://doi.org/10.1021/ci00028a014

Hunt, C. M., Westerkam, W. R., & Stave, G. M. (1992). Effect of age and gender on the activity of human hepatic CYP3A. *Biochemical Pharmacology*, *44*(2), 275–283. https://doi.org/10.1016/0006-2952(92)90010-G

Ibanez, L., Perez, E., Vidal, X., Laporte, J. R., & Grup d'Estudi Multicenteric d'Hepatotoxicitat Aguda de, B (2002). Prospective surveillance of acute serious liver disease unrelated to infectious, obstructive, or metabolic diseases: epidemiological and clinical features, and exposure to drugs. *Journal of Hepatology*, *37*(5), 592–600.

Jornil, J., Nielsen, T. S., Rosendal, I., Ahlner, J., Zackrisson, A. L., Boel, L. W., & Brock, B. (2013). A poor metabolizer of both CYP2C19 and CYP2D6 identified by mechanistic pharmacokinetic simulation in a fatal drug poisoning case involving venlafaxine. *Forensic Science International*, *226*(1–3), e26–e31. https://doi.org/10.1016/j.forsciint.2012.12.020

Kock, K., Ferslew, B. C., Netterberg, I., Yang, K., Urban, T. J., Swaan, P. W., … Brouwer, K. L. (2014). Risk factors for development of cholestatic drug-induced liver injury: inhibition of hepatic basolateral bile acid transporters multidrug resistance-associated proteins 3 and 4. *Drug*

Metabolism and Disposition, *42*(4), 665–674. https://doi.org/10.1124/dmd.113.054304

Lammert, C., Einarsson, S., Saha, C., Niklasson, A., Bjornsson, E., & Chalasani, N. (2008). Relationship between daily dose of oral medications and idiosyncratic drug-induced liver injury: search for signals. *Hepatology*, *47*(6), 2003–2009. https://doi.org/10.1002/hep.22272

Lasser, K. E., Allen, P. D., Woolhandler, S. J., Himmelstein, D. U., Wolfe, S. M., & Bor, D. H. (2002). Timing of new black box warnings and withdrawals for prescription medications. *JAMA*, *287*(17), 2215–2220. https://doi.org/10.1001/jama.287.17.2215

Li, X., & Cameron, M. D. (2012). Potential role of a quetiapine metabolite in quetiapine-induced neutropenia and agranulocytosis. *Chemical Research in Toxicology*, *25*(5), 1004–1011. https://doi.org/10.1021/tx2005635

Lima, A., Bernardes, M., Azevedo, R., Medeiros, R., & Seabra, V. (2015). Pharmacogenomics of methotrexate membrane transport pathway: Can clinical response to methotrexate in rheumatoid arthritis be predicted? *International Journal of Molecular Sciences*, *16*(6), 13760–13780. https://doi.org/10.3390/ijms160613760

Makhlouf, H. A., Helmy, A., Fawzy, E., El-Attar, M., & Rashed, H. A. (2008). A prospective study of antituberculous drug-induced hepatotoxicity in an area endemic for liver diseases. *Hepatology International*, *2*(3), 353–360. https://doi.org/10.1007/s12072-008-9085-y

Miners, J. O., & Birkett, D. J. (1998). Cytochrome P4502C9: an enzyme of major importance in human drug metabolism. *British Journal of Clinical Pharmacology*, *45*(6), 525–538. https://doi.org/10.1046/j.1365-2125.1998.00721.x

Morgan, R. E., Trauner, M., van Staden, C. J., Lee, P. H., Ramachandran, B., Eschenberg, M., … Hamadeh, H. K. (2010). Interference with bile salt export pump function is a susceptibility factor for human liver injury in drug development. *Toxicological Sciences*, *118*(2), 485–500. https://doi.org/10.1093/toxsci/kfq269

Najibi, A., Heidari, R., Zarifi, J., Jamshidzadeh, A., Firoozabadi, N., & Niknahad, H. (2016). Evaluating the role of drug metabolism and reactive intermediates in trazodone-induced cytotoxicity toward freshly-isolated rat hepatocytes. *Drug Research*, *66*(11), 592–596. https://doi.org/10.1055/s-0042-109536

Niederer, C., Behra, R., Harder, A., Schwarzenbach, R. P., & Escher, B. I. (2004). Mechanistic approaches for evaluating the toxicity of reactive organochlorines and epoxides in green algae. *Environmental Toxicology and Chemistry*, *23*(3), 697–704. https://doi.org/10.1897/03-83

Parkinson, A., Mudra, D. R., Johnson, C., Dwyer, A., & Carroll, K. M. (2004). The effects of gender, age, ethnicity, and liver cirrhosis on cytochrome P450 enzyme activity in human liver microsomes and inducibility in cultured human hepatocytes. *Toxicology and Applied Pharmacology*, *199*(3), 193–209. https://doi.org/10.1016/j.taap.2004.01.010

Pellinen, P., Honkakoski, P., Stenback, F., Niemitz, M., Alhava, E., Pelkonen, O., … Pasanen, M. (1994). Cocaine N-demethylation and the metabolism-related hepatotoxicity can be prevented by cytochrome P450 3A inhibitors. *European Journal of Pharmacology*, *270*(1), 35–43.

Rendic, S. (2002). Summary of information on human CYP enzymes: human P450 metabolism data. *Drug Metabolism Reviews*, *34*(1–2), 83–448. https://doi.org/10.1081/DMR-120001392

Reuben, A., Koch, D. G., Lee, W. M., & Acute Liver Failure Study, G. (2010). Drug-induced acute liver failure: results of a U.S. multicenter, prospective study. *Hepatology*, *52*(6), 2065–2076. https://doi.org/10.1002/hep.23937

Rodrigues, A. C. (2010). Efflux and uptake transporters as determinants of statin response. *Expert Opinion on Drug Metabolism & Toxicology*, *6*(5), 621–632. https://doi.org/10.1517/17425251003713519

Rücker, C., Rücker, G., & Meringer, M. (2007). y-Randomization and its variants in QSPR/QSAR. *Journal of Chemical Information and Modeling*, *47*(6), 2345–2357. https://doi.org/10.1021/ci700157b

Stine, J. G., & Chalasani, N. P. (2017). Drug hepatotoxicity: Environmental factors. *Clinical Liver Disease*, *21*(1), 103–113. https://doi.org/10.1016/j.cld.2016.08.008

Stine, J. G., & Lewis, J. H. (2011). Drug-induced liver injury: a summary of recent advances. *Expert Opinion on Drug Metabolism & Toxicology*, *7*(7), 875–890. https://doi.org/10.1517/17425255.2011.577415

Teh, L. K., & Bertilsson, L. (2012). Pharmacogenomics of CYP2D6: molecular genetics, interethnic differences and clinical importance. *Drug Metabolism and Pharmacokinetics*, *27*(1), 55–67. https://doi.org/10.2133/dmpk.DMPK-11-RV-121

Thompson, R. A., Isin, E. M., Ogese, M. O., Mettetal, J. T., & Williams, D. P. (2016). Reactive metabolites: Current and emerging risk and hazard assessments. *Chemical Research in Toxicology*, *29*(4), 505–533. https://doi.org/10.1021/acs.chemrestox.5b00410

Walker, K., Ginsberg, G., Hattis, D., Johns, D. O., Guyton, K. Z., & Sonawane, B. (2009). Genetic polymorphism in N-acetyltransferase (NAT): Population distribution of NAT1 and NAT2 activity. *Journal of Toxicology and Environmental Health. Part B, Critical Reviews*, *12*(5–6), 440–472. https://doi.org/10.1080/10937400903158383

Yu, K., Geng, X., Chen, M., Zhang, J., Wang, B., Ilic, K., & Tong, W. (2014). High daily dose and being a substrate of cytochrome P450 enzymes are two important predictors of drug-induced liver injury. *Drug Metabolism and Disposition*, *42*(4), 744–750. https://doi.org/10.1124/dmd.113.056267

Zanger, U. M., Turpeinen, M., Klein, K., & Schwab, M. (2008). Functional pharmacogenetics/genomics of human cytochromes P450 involved in drug biotransformation. *Analytical and Bioanalytical Chemistry*, *392*(6), 1093–1108. https://doi.org/10.1007/s00216-008-2291-6

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Hammann F, Schöning V, Drewe J. Prediction of clinically relevant drug-induced liver injury from structure using machine learning. *J Appl Toxicol*. 2019;39:412–419. https://doi.org/10.1002/jat.3741