



Vienna LiverTox Workspace—A Set of Machine Learning Models for Prediction of Interactions Profiles of Small Molecules With Transporters Relevant for Regulatory Agencies

Floriane Montanari[†], Bernhard Knasmüller, Stefan Kohlbacher, Christoph Hillisch, Christine Baierová, Melanie Grandits* and Gerhard F. Ecker

OPEN ACCESS

Pharmacoinformatics Research Group, Department of Pharmaceutical Chemistry, University of Vienna, Vienna, Austria

Edited by:

Jose L. Medina-Franco,
National Autonomous University of
Mexico, Mexico

Reviewed by:

Vinicius M. Alves,
University of North Carolina at Chapel
Hill, United States
Rodolpho C. Braga,
InsilicAll, Brazil

*Correspondence:

Melanie Grandits
melanie.grandits@univie.ac.at

[†]Present address:

Floriane Montanari,
Department of Digital Technologies,
Bayer AG, Berlin, Germany

Specialty section:

This article was submitted to
Medicinal and Pharmaceutical
Chemistry,
a section of the journal
Frontiers in Chemistry

Received: 09 October 2019

Accepted: 13 December 2019

Published: 10 January 2020

Citation:

Montanari F, Knasmüller B,
Kohlbacher S, Hillisch C, Baierová C,
Grandits M and Ecker GF (2020)
Vienna LiverTox Workspace—A Set of
Machine Learning Models for
Prediction of Interactions Profiles of
Small Molecules With Transporters
Relevant for Regulatory Agencies.
Front. Chem. 7:899.
doi: 10.3389/fchem.2019.00899

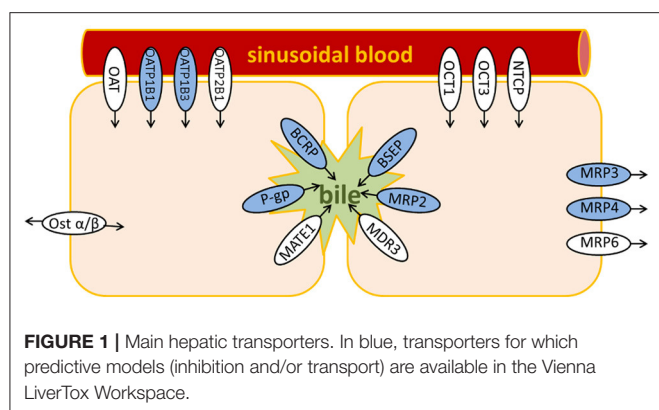
Transporters expressed in the liver play a major role in drug pharmacokinetics and are a key component of the physiological bile flow. Inhibition of these transporters may lead to drug-drug interactions or even drug-induced liver injury. Therefore, predicting the interaction profile of small molecules with transporters expressed in the liver may help medicinal chemists and toxicologists to prioritize compounds in an early phase of the drug development process. Based on a comprehensive analysis of the data available in the public domain, we developed a set of classification models which allow to predict—for a small molecule—the inhibition of and transport by a set of liver transporters considered to be relevant by FDA, EMA, and the Japanese regulatory agency. The models were validated by cross-validation and external test sets and comprise cross validated balanced accuracies in the range of 0.64–0.88. Finally, models were implemented as an easy to use web-service which is freely available at <https://livertox.univie.ac.at>.

Keywords: Vienna LiverTox Workspace, web service, machine learning, ABC-transporter, OATP-transporter, toxicity, classification models

INTRODUCTION

Membrane transporters expressed in the liver play different, but interconnected roles: on the one hand, basolateral transporters pick up xenobiotics, and endogenous molecules from the portal vein to the liver, or excrete their substrates into the blood. Apical transporters, on the other hand, take care of the flux toward the bile duct network (**Figure 1**). Three main types of substrates are of interest with respect to liver toxicity: drugs, which enter the hepatocytes at the first hepatic pass or at the elimination stage; bilirubin, a product of the degradation of the heme; and bile salts, which circulate between the gastro-intestinal tract, and the liver.

Additionally, to the enzyme family of cytochromes, also the transporters expressed in the liver are crucial for a fully functional organ. Some of them are e.g., involved in the bilirubin cycle: OATP1B1 and OATP1B3 uptake bilirubin into the hepatocytes (Briz et al., 2003), where glucuronidation takes place. MRP2 then excretes the bilirubin conjugate to the bile (Kamisako et al., 1999). At the basolateral membrane, MRP3 might also excrete it back to the sinusoidal blood (Keppler, 2014). As a result, inhibition of the uptake OATP transporters or of MRP2 may



lead to the accumulation of bilirubin (conjugated or not) in the blood, which is referred to as hyperbilirubinemia. Conjugated hyperbilirubinemia is a marker of hepatobiliary injury (Dufour et al., 2000; Ozer et al., 2008; Padda et al., 2011), and predicting it may allow to flag compounds that could cause liver injury.

Bile acids are synthesized in the liver by catabolism of cholesterol and then excreted to the bile by the active bile salt export pump (BSEP) and by the multidrug resistance-associated protein 2 (MRP2) (Meier and Stieger, 2002). Bile salts have a pronounced detergent effect, which explains their toxicity when they accumulate in the liver (Attili et al., 1986). For their transport in the bile duct, bile salts form mixed micelles with phospholipids of the outer leaflet of the membrane. The multidrug resistance protein 3 (MDR3) allows lipid flopping at the apical membrane of the hepatocyte, and its function is necessary to avoid bile duct toxicity (Nicolau et al., 2012). After reaching the intestine via the bile flow, bile acids are reabsorbed into the portal vein, and taken up again into the hepatocytes by the sodium taurocholate co-transporting polypeptide (NTCP) (Stieger, 2011). Impairment of bile flow leading to a toxic accumulation of bile salts in the hepatocytes might lead to drug-induced cholestasis, which is one of the main causes of drug-induced liver injury (DILI) (Padda et al., 2011).

Apart from playing a role in proper bile flow and bilirubin elimination, liver transporters also transport drugs that will then be metabolized and excreted. At this stage, drugs can inhibit different transporters and cause drug-drug interactions (König et al., 2013) (in case of co-administered inhibitor and substrate) or liver injury (by disrupting the bile flow for example). This is why predicting the inhibition and the substrate profile for liver transporters might be useful in identifying potentially problematic compounds. In addition, the Food and Drugs administration recommends experimental testing of the

Abbreviations: ADME-Tox, absorption distribution metabolism excretion toxicity; BCRP, breast cancer resistance protein; BSEP, bile salt export pump; DILI, drug-induced liver injury; ECFP, extended connectivity fingerprint; MDR1, multidrug resistance protein 1; MDR3, multidrug resistance protein 3; MRP2, multidrug resistance-associated protein 2; MRP3, multidrug resistance-associated protein 3; MRP4, multidrug resistance-associated protein 4; NTCP, sodium taurocholate co-transporting polypeptide; OATP1B1, organic anion transporting polypeptide 1B1; OATP1B3, organic anion transporting polypeptide 1B3; P-gp, P-glycoprotein; SMILES, simplified molecular input line entry system.

interactions between drugs and transporters (especially P-gp, BCRP, OATP1B1, and OATP1B3) to identify potential drug-drug interactions (U.S. Department of Health Human Services Food Drug Administration Center for Drug Evaluation Research, 2012). Thus, it definitely would be of value to have a suite of computational models available which allow the fast and easy assessment of compounds for their interaction profile with transporters expressed in the liver.

Here, we present the Vienna LiverTox Workspace, a web server for the prediction of interactions with liver transporters as well as selected liver toxicity endpoints. To the best of our knowledge, it is the first time that such an ensemble of predictive models for hepatotoxicity and liver transport is made available to the public. The predictions are made by individual machine learning models built on publicly available data for each target of interest.

METHODS

Data Curation

The datasets for the training and testing of the models were collected from different sources (online tools as well as publications). The data were cleaned by using an in-house system combining the molecular operating environment (MOE 2014.09) (Molecular Operating Environment, 2014) wash option and the Atkinson Standardiser (<https://github.com/flatkinson/standardiser>). This approach was used for some of the datasets and for others the cleaning procedure of already published papers were used (Pinto et al., 2012; Kotsampasakou et al., 2015). In general, duplicates were removed from the dataset, including pairs of stereoisomers. Further, if these compounds share the same class label, one of the compounds was kept. A detailed list of references as well as the number of compounds revealed after the preprocessing of the data is given in **Table S1**. In the sections Transporter Models and Hepatotoxicity Models, the data curation as well as the model generation for the specific endpoints is given. The datasets are available in the Supplementary Material (**Data Sheet 2**).

Transporter Models

The web service allows for the prediction of interactions between a small molecule and eight different liver transporters (**Figure 1**, transporters marked in blue). The lack of publicly available data for the other transporters explains the absence of respective models in the Vienna LiverTox Workspace.

All the models predicting whether a compound will be a substrate of a transporter (BCRP, P-gp, BSEP, MRP2, and MRP3) were built upon a dataset correlating expression levels of 47 ABC-transporters with drug toxicity, which can then be used to infer transported vs. non-transported compounds (Szakács et al., 2004). For the transport inhibition models (BCRP, P-gp, BSEP, MRP3, MRP4, OATP1B1, and OATP1B3) the datasets were collected from literature and, if necessary, manually aggregated. In both cases, the models predict a binary outcome: the query compound is a substrate or not, or an inhibitor or not.

For chemistry encoding of the compounds, we used circular fingerprints or 2D molecular descriptors as implemented in

TABLE 1 | Overview of the parameters used for the models.

	Descriptors	Trainings set	Algorithm
INHIBITION			
P-gp (MDR1)	ECFP8-like fingerprints	Broccatelli et al., 2011	SVM
BSEP	Molecular descriptors	Warner et al., 2012 Dawson et al., 2012 Morgan et al., 2010	Naïve bayes
BCRP	ECFP8-like fingerprints	Montanari and Ecker, 2014	Logistic regression
MRP3	Molecular descriptors	Köck et al., 2014	BayesNet
MRP4	Molecular descriptors	Köck et al., 2014	AdaBoost (MetaCost)
OATP1B1	Molecular descriptors	De Bruyn et al., 2013	BayesNet
OATP1B3	Molecular descriptors	De Bruyn et al., 2013	BayesNet
TRANSPORT			
P-gp (MDR1)	Molecular descriptors	Szakács et al., 2004	Rotation Forest (MetaCost)
BSEP			SVM (MetaCost)
BCRP			k-nearest neighbors (MetaCost)
MRP2			
MRP3			
TOXICITY			
Hyperbilirubinemia	ECFP8-like fingerprints	Liu et al., 2011	SVM (MetaCost)
Cholestasis	Molecular descriptors	SIDER v2 database (Kuhn et al., 2010, 2016)	Tree model (MetaCost)
Drug-induced liver injury (DILI)	Molecular descriptors	Various sources*	Random Forest

*see <https://livertox.univie.ac.at/> for a detailed list.

RDKit version 2015.03.1 (<https://www.rdkit.org/>). Different machine learning algorithms were applied and the one giving the best cross-validation results was kept as final model. Especially for the transport prediction, a heavy class imbalance (most of the drugs in the training set were non-substrates) was noted, which was handled by MetaCost (Domingos, 1999). The exact methodology and cross-validation performance for each individual transporter model is described in the documentation available at <https://livertox.univie.ac.at>, and an overview is given in **Table 1**. In some cases, external test sets were collected from Metrabase (Mak et al., 2015) or from recent publications (time-split evaluation).

Hepatotoxicity Models

Three models in the web service can be used to assess human liver damage potentially caused by a test compound: hyperbilirubinemia, cholestasis, and drug-induced liver injury (DILI).

For hyperbilirubinemia, 835 compounds were taken from Kotsampasakou et al. (2017a) and the modeling methodology was kept as in Kotsampasakou et al. (2017b): ECFP-like fingerprints were computed with RDKit (<https://www.rdkit.org/>), then a combination of feature selection, MetaCost, and support vector machines with RBF kernel was used for learning. The cholestasis model uses data from Kotsampasakou and Ecker (2017) and a combination of MetaCost and a tree algorithm to predict whether a compound is likely to cause cholestasis or not. Finally, the DILI model is based on a 966-compound dataset carefully compiled from literature (Kotsampasakou et al., 2017c). RDKit molecular descriptors and a random forest of 500 trees are used for modeling.

Web Service Implementation

The Vienna LiverTox Workspace has been implemented as Python/php based web service. It consists of two parts, namely the backend and the frontend. The backend consists of a docker image which runs the machine learning models on an input SD-File. It consists of a Python Flask server (<https://palletsprojects.com/p/flask/>) which processes the requests from the frontend. Each request consists of one or more input molecules and a list of models to run the predictions on. The frontend, also a docker image, is based on the CakePHP framework (<https://cakephp.org/>) and is responsible for the user interface (UI), which sends the request to the backend and displays the results. The web service provides, after a login, the possibility to upload a SD-File of 10 compounds. The service can also be used without logging in but then it is only possible to draw and predict a single molecule. JSME (Bienfait and Ertl, 2013) is used as drawing tool. The web service runs on an Ubuntu Linux based server with two twelve-core Intel Xeon 64bit processors and 128 GB RAM, and is hosted at the University of Vienna by the Pharmacoinformatics Research Group.

The models use the RDKit library (<https://www.rdkit.org/>) (version 2015.03.1) for computing the descriptors and handling the chemistry aspects, while the Weka (Hall et al., 2009) (version 3.7.11) and scikit-learn (Pedregosa et al., 2011) (version 0.14.1) libraries are used to train and run the predictive models. The models also include a compound cleaning step, implemented with the Atkinson Standardiser (<https://github.com/flatkinson/standardiser>) (**Figure 2**).

The steps performed during model building and a test compound prediction are shown in **Figure 2**. In both cases, the compound is standardized and the molecular descriptors are

calculated. In the case of the model generation, this allows the training of the model and its development (left workflow). For the prediction of a test compound, the descriptors are passed to the available model to predict its class affiliation (right workflow).

In general, the output of the model gives, in addition to the class prediction, the actual score. This score is a numerical value between 0 and 1, and roughly corresponds to a probability of being active (inhibitor, substrate, or toxic compound). Therefore, a value close to 1 indicates substrate/inhibitor/toxic properties, a value close to 0 annotates for non-substrates/non-inhibitors/non-toxic.

Applicability Domain

The Applicability Domain (AD) is used to validate the reliability of a given prediction model. It defines whether

a dataset of interest is in or out of domain, meaning if it falls within the chemical space of the model or not. If it is out of Domain, the prediction cannot be regarded as reliable.

In our study, an Applicability Domain model, using the approach of Sahigara et al. (2013), was created for each transporter with the respective training set. RDKit descriptors were used as molecule representation. This approach combines the classical, widely used k-Nearest Neighbor (k-NN) method with a probability density function estimation. It uses three stages to determine the reliable space of a prediction model. First, a set of thresholds is defined depending on the diverse densities of the training set by considering the 15 k-nearest neighbors using Euclidean distances. This allows the AD to consider a dense and sparse training region (The threshold defines if a test sample can be reliably predicted). In a next step a decision rule is derived to filter out outlier molecules. Finally, the reliability of the AD is tested by looking at the model statistics and prediction errors. This feature is not yet implemented in the Web service, but will be available soon.

RESULTS

Model Results

The performance of the models was estimated by calculating statistical performance metrics using a 10-fold cross-validation. The results are provided in Table 2. The overall accuracy, corresponding to the rate of correct predictions, ranges from 0.59 to 0.87. Also, the sensitivity of the models was calculated (0.57–0.85). This parameter gives the number of actual positives that are correctly identified and is expressed by the number of true positives divided by the number of positive predictions. Further, the number of actual negatives was determined by the

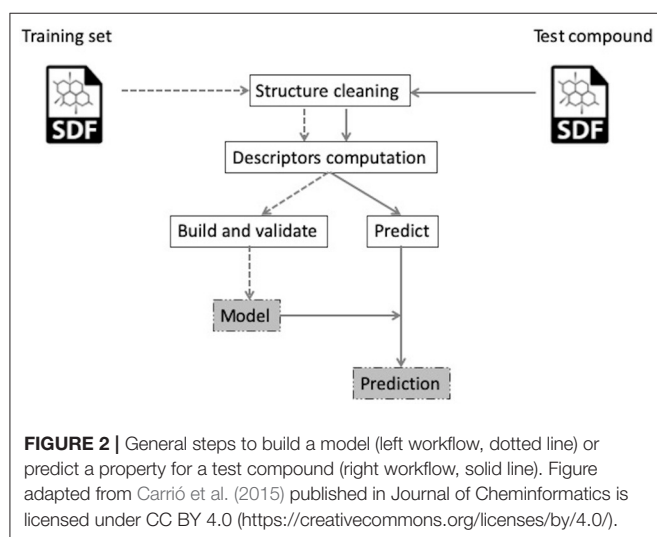


TABLE 2 | Performance metrics of the transporter models.

	Accuracy	Sensitivity	Specificity	Balanced accuracy	MCC	ROC AUC
INHIBITION						
P-gp (MDR1)	0.87	0.85	0.90	0.88	0.76	0.94
BSEP	0.85	0.77	0.87	0.82	0.60	0.91
BCRP	0.83	0.77	0.87	0.82	0.64	0.90
MRP3	0.82	0.75	0.87	0.81	0.62	0.86
MRP4	0.73	0.86	0.65	0.76	0.49	0.74
OATP1B1	0.77	0.71	0.78	0.75	0.34	0.80
OATP1B3	0.81	0.78	0.81	0.80	0.35	0.81
TRANSPORT						
P-gp (MDR1)	0.81	0.81	0.81	0.81	0.44	0.85
BSEP	0.71	0.65	0.72	0.69	0.28	0.69
BCRP	0.73	0.57	0.75	0.66	0.20	0.71
MRP2	0.70	0.60	0.72	0.66	0.27	0.73
MRP3	0.72	0.60	0.74	0.67	0.28	0.72
TOXICITY						
Hyperbilirubinemia	0.67	0.64	0.67	0.66	0.20	0.69
Cholestasis	0.59	0.72	0.56	0.64	0.22	0.64
Drug-induced liver injury (DILI)	0.65	0.72	0.58	0.65	0.30	0.70

Select Models

SELECT ALL

SELECT TRANSPORTER MODELS:

BSEP Inhibition

BSEP Transport

P-glycoprotein Inhibition

P-glycoprotein Transport

MRP4 Inhibition

MRP2 Transport

MRP3 Inhibition

MRP3 Transport

BCRP Inhibition

BCRP Transport

OATP1B1 Inhibition

OATP1B3 Inhibition

SELECT TOXICITY MODELS:

Drug-induced liver injury

Hyperbilirubinemia

Cholestasis

Draw Compound

O=C(CCc1ccccc1)c2ccccc2OCCN4CC(=O)N3CCCC3C4=O

Results

Model	Classification	Score
BSEP Inhibition	positive	1
BSEP Transport	negative	0
P-glycoprotein Inhibition	positive	0.93
P-glycoprotein Transport	negative	0.01
MRP4 Inhibition	positive	1
MRP2 Transport	positive	0.6
MRP3 Inhibition	positive	1
MRP3 Transport	positive	0.8
BCRP Inhibition	negative	0.4
BCRP Transport	negative	0.4
OATP1B1 Inhibition	positive	1
OATP1B3 Inhibition	positive	1
Drug-induced liver injury	negative	0.45
Hyperbilirubinemia	positive	0.69
Cholestasis	positive	0.8

Want more?
If you share our love for fully-integrated toxicology systems, do not hesitate to contact us: gerhard.f.ecker@univie.ac.at

FIGURE 3 | Overview of the web service interface. On the left side, test compound is drawn and desired models are selected. On the right side, results table with the predictions and model scores.

number of true negatives divided by the number of negative predictions. The so-called specificity ranges from 0.56 to 0.90. To estimate a metric for the quality of the models, the Matthews correlation coefficient (MCC) and the Area under the Receiver Operating Characteristics curve (ROC AUC) were determined. The MCC is a number between -1 and 1 where 0 indicates a prediction equal to a random prediction and 1 indicates a perfect prediction, whereas -1 is a complete miss. The scores for our models are in the range of 0.20 – 0.76 . The ROC AUC measures the ability of the model to distinguish between negatives and positives, while a higher value indicates a better performance. In the models provided on the Vienna LiverTox Workspace the ROC AUC ranges from 0.64 to 0.94 . Furthermore, if data was available, the models were also validated with one or more external test sets. For more details see the documentation on the website.

Use Case: Prediction of Liver Interaction for a Propafenone Analog

In this section, we briefly detail how predictions can be generated for a given compound using the LiverTox web service. In first instance, the compound is drawn or its SMILES string is pasted in the Molecule Editor (Bienfait and Ertl, 2013). Then the models can be selected on the left panel, either one by one or all at the same time (Figure 3, left side).

By clicking on the SUBMIT button, the data is sent to the backend server where the predictions are running. Upon completion of the calculations, a table listing the different models and corresponding outputs will appear (Figure 3, right side). The second column in the results table corresponds to the binary classification, while the third column “Score” gives the actual output of the model, which corresponds to a probability of being active. For example, for BCRP inhibition and transport

TABLE 3 | Comparison of existing free online tools to predict ADME-Tox properties of compounds.

Web service	Transporters predictions	CYP450 predictions	Hepatotox. predictions	Batch prediction	Run time for 1 compound
ProTox-II (Drwal et al., 2014)	No	No	No	Yes (max. 100)	<5 s
BioZyne (Levatić et al., 2013)	P-gp	No	No	Not for free	~5 s
QSAR DB (http://qsar.food.dtu.dk/)	No	Yes	No	Yes	N.A.
pkCSM (Pires et al., 2015)	P-gp	Yes	Yes	Yes (max. 100)	<5 s for 30 models
Lazar (Maunz et al., 2013)	No	No	No	No	~10 s for 6 models
Vienna LiverTox Workspace	P-gp, BSEP, BCRP, MRP2, MRP3, MRP4, OATP1B1, OATP1B3	No	Yes	Not for free	~30 s for 15 models

and for DILI, the output score is close to 0.5 (which is the threshold used to separate predicted actives from predicted inactives), which indicates an uncertainty of these three models for the particular query compound. Propafenone derivatives are frequently reported as inhibitors of P-gp, and indeed the P-gp inhibition model predicts this particular one to be an inhibitor with a high score (0.93).

DISCUSSION

Many systems already exist to predict *in silico* activities or properties of small molecules. **Table 3** compares freely available ones with our own web service in terms of model offer, submission and run time. For example, ProTox-II predicts oral drug toxicity in rodents (lethal dose LD₅₀ and a category of toxicity between 1 and 6) using similarity to compounds with known LD₅₀ and recognition of toxic fragments (Drwal et al., 2014). BioZyne proposes exclusively one model for P-gp transport prediction based on the same dataset as ours (Szakács et al., 2004; Levatić et al., 2013). It uses a Support Vector Machine classifier for the prediction of P-gp substrates. The Danish (Q)SAR Database contains pre-calculated properties combined from more than 200 models from both commercial and free tools (<http://qsar.food.dtu.dk/>). Predictions for environmental toxicity, blood-brain barrier permeation, cytochrome interactions, or human genotoxicity are available. Unfortunately, new predictions for compounds that are not part of the database cannot be made. PkCSM is another web service for predicting pharmacokinetics properties of compounds (Pires et al., 2015). Models such as P-gp inhibition and transport, blood-brain barrier permeation, interaction with cytochromes, renal clearance, or even liver toxicity are available.

In general, our models for the inhibitors show a better performance especially when looking at the correct prediction of the positives. The prediction of true negatives is for the inhibitor and transporter models quite similar which can be explained by the availability of more negatives if the training set is unbalanced. This is especially the case for the substrate models. The quality of the prediction (MCC) is higher for the inhibition models of P-gp, BSEP, BCRP, and MRP3 since the available dataset is more balanced. In comparison, the three toxicity models show a poorer performance due to the complexity of these endpoints and especially for hyperbilirubinemia and cholestasis which shows also a lack of positives.

The Transporters selected for this web service were chosen based on their importance for regulatory agencies such as FDA, EMA and the Japanese regulatory agency. They recommend or in some cases request these proteins to be routinely tested in inhibition—and substrate studies of new drugs.

CONCLUSION

We have presented the Vienna LiverTox Workspace, a web service dedicated to the prediction of liver toxicity and interactions between small molecules and liver transporters. It is easy to use, fast, web browser agnostic, and well-documented. Thanks to its modular system, it will be easy to integrate new models in the future, as well as re-implement existing models in case new training data becomes available. We hope that our models will help researchers to flag potentially dangerous compounds and shed light on the relationships between liver transporters and toxicity.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

FM, SK, CH, and MG developed the models. BK, CB, and MG implemented the web service. MG and GE supervised the study. FM and MG wrote the majority of the manuscript. All authors contributed to refining the manuscript.

FUNDING

We gratefully acknowledge financial support provided by the Austrian Science Fund, grant #F03502 (SFB35), and by the Innovative Medicines Initiative Joint Undertaking under grant agreement n°115002 (eTOX).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fchem.2019.00899/full#supplementary-material>

REFERENCES

- Attili, A. F., Angelico, M., Cantafora, A., Alvaro, D., and Capocaccia, L. (1986). Bile acid-induced liver toxicity: relation to the hydrophobic-hydrophilic balance of bile acids. *Med. Hypotheses* 19, 57–69. doi: 10.1016/0306-9877(86)90137-4
- Bienfait, B., and Ertl, P. (2013). JSME: a free molecule editor in JavaScript. *J. Cheminform.* 5:24. doi: 10.1186/1758-2946-5-24
- Briz, O., Serrano, M. A., Macías, R. I. R., Gonzalez-Gallego, J., and Marin, J. J. G. (2003). Role of organic anion-transporting polypeptides, OATP-A, OATP-C and OATP-8, in the human placenta-maternal liver tandem excretory pathway for foetal bilirubin. *Biochem. J.* 371, 897–905. doi: 10.1042/bj20030034
- Broccatelli, F., Carosati, E., Neri, A., Frosini, M., Goracci, L., Oprea, T. I., et al. (2011). A novel approach for predicting P-glycoprotein (ABCB1) inhibition using molecular interaction fields. *J. Med. Chem.* 54, 1740–1751. doi: 10.1021/jm101421d
- Carrió, P., López, O., Sanz, F., and Pastor, M. (2015). eTOXlab, an open source modeling framework for implementing predictive models in production environments. *J. Cheminform.* 7:8. doi: 10.1186/s13321-015-0058-6
- Dawson, S., Stahl, S., Paul, N., Barber, J., and Kenna, J. G. (2012). *In vitro* inhibition of the bile salt export pump correlates with risk of cholestatic drug-induced liver injury in humans. *Drug Metab. Dispos.* 40, 130–8. doi: 10.1124/dmd.111.040758
- De Bruyn, T., van Westen, G. J. P., IJzerman, A. P., Stieger, B., de Witte, P., Augustijns, P. F., et al. (2013). Structure-based identification of OATP1B1/3 inhibitors. *Mol. Pharmacol.* 83, 1257–1267. doi: 10.1124/mol.112.084152
- Domingos, P. (1999). “MetaCost: a general method for making classifiers cost-sensitive,” in *Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (New York), 155–164. doi: 10.1145/312129.312220
- Drwal, M. N., Banerjee, P., Dunkel, M., Wettig, M. R., and Preissner, R. (2014). ProTox: a web server for the *in silico* prediction of rodent oral toxicity. *Nucleic Acids Res.* 42, W53–W58. doi: 10.1093/nar/gku401
- Dufour, D. R., Lott, J. A., Nolte, F. S., Gretch, D. R., Koff, R. S., and Seeff, L. B. (2000). Diagnosis and monitoring of hepatic injury. II. Recommendations for use of laboratory tests in screening, diagnosis, and monitoring. *Clin. Chem.* 46, 2050–68.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. (2009). The WEKA data mining software: an update. *SIGKDD Explor. Newsl.* 11, 10–18. doi: 10.1145/1656274.1656278
- Kamisako, T., Leier, I., Cui, Y., König, J., Buchholz, U., Hummel-Eisenbeiss, J., et al. (1999). Transport of monoglucuronosyl and bisglucuronosyl bilirubin by recombinant human and rat multidrug resistance protein 2. *Hepatology* 30, 485–490. doi: 10.1002/hep.510300220
- Keppeler, D. (2014). The roles of MRP2, MRP3, OATP1B1, and OATP1B3 in conjugated hyperbilirubinemia. *Drug Metab. Dispos.* 42, 561–565. doi: 10.1124/dmd.113.055772
- Köck, K., Ferslew, B. C., Netterberg, I., Yang, K., Urban, T. J., Swaan, P. W., et al. (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 Metab. Dispos.* 42, 665–74. doi: 10.1124/dmd.113.054304
- König, J., Müller, F., and Fromm, M. F. (2013). Transporters and drug-drug interactions: important determinants of drug disposition and effects. *Pharmacol. Rev.* 65, 944–966. doi: 10.1124/pr.113.007518
- Kotsampasakou, E., Brenner, S., Jäger, W., and Ecker, G. F. (2015). Identification of novel inhibitors of organic anion transporting polypeptides 1B1 and 1B3 (OATP1B1 and OATP1B3) using a consensus vote of six classification models. *Mol. Pharm.* 12, 4395–4404. doi: 10.1021/acs.molpharmaceut.5b00583
- Kotsampasakou, E., and Ecker, G. F. (2017). Predicting drug-induced cholestasis with the help of hepatic transporters—an *in silico* modeling approach. *J. Chem. Inf. Model.* 57, 608–615. doi: 10.1021/acs.jcim.6b00518
- Kotsampasakou, E., Escher, S. E., and Ecker, G. F. (2017a). Curated human hyperbilirubinemia data and the respective OATP1B1 and 1B3 inhibition predictions. *Data Br.* 11, 204–207. doi: 10.1016/j.dib.2017.02.009
- Kotsampasakou, E., Escher, S. E., and Ecker, G. F. (2017b). Linking organic anion transporting polypeptide 1B1 and 1B3 (OATP1B1 and OATP1B3) interaction profiles to hepatotoxicity—The hyperbilirubinemia use case. *Eur. J. Pharm. Sci.* 100, 9–16. doi: 10.1016/j.ejps.2017.01.002
- Kotsampasakou, E., Montanari, F., and Ecker, G. F. (2017c). Predicting drug-induced liver injury: the importance of data curation. *Toxicology* 389, 139–145. doi: 10.1016/j.tox.2017.06.003
- Kuhn, M., Campillos, M., Letunic, I., Jensen, L. J., and Bork, P. (2010). A side effect resource to capture phenotypic effects of drugs. *Mol. Syst. Biol.* 6:343. doi: 10.1038/msb.2009.98
- Kuhn, M., Letunic, I., Jensen, L. J., and Bork, P. (2016). The SIDER database of drugs and side effects. *Nucleic Acids Res.* 44, D1075–D1079. doi: 10.1093/nar/gkv1075
- Levatić, J., Curak, J., Kralj, M., Šmuc, T., Osmak, M., and Supek, F. (2013). Accurate models for P-gp drug recognition induced from a cancer cell line cytotoxicity screen. *J. Med. Chem.* 56, 5691–5708. doi: 10.1021/jm400328s
- Liu, Z., Shi, Q., Ding, D., Kelly, R., Fang, H., and Tong, W. (2011). Translating clinical findings into knowledge in drug safety evaluation - drug induced liver injury prediction system (DILLps). *PLoS Comput. Biol.* 7:e1002310. doi: 10.1371/journal.pcbi.1002310
- Mak, L., Marcus, D., Howlett, A., Yarova, G., Duchateau, G., Klaffke, W., et al. (2015). Metrabase: a cheminformatics and bioinformatics database for small molecule transporter data analysis and (QSAR) modeling. *J. Cheminform.* 7:31. doi: 10.1186/s13321-015-0083-5
- Maunz, A., Gütlein, M., Rautenberg, M., Vorgrimmler, D., Gebele, D., and Helma, C. (2013). Lazar: a modular predictive toxicology framework. *Front. Pharmacol.* 4:38. doi: 10.3389/fphar.2013.00038
- Meier, P. J., and Stieger, B. (2002). Bile salt transporters. *Annu. Rev. Physiol.* 64, 635–661. doi: 10.1146/annurev.physiol.64.082201.100300
- Molecular Operating Environment (2014). *Molecular Operating Environment*. Available at: <https://www.chemcomp.com/>
- Montanari, F., and Ecker, G. F. (2014). BCRP inhibition: from data collection to ligand-based modeling. *Mol. Inform.* 33, 322–331. doi: 10.1002/minf.201400012
- Morgan, R. E., Trauner, M., van Staden, C. J., Lee, P. H., Ramachandran, B., Eschenberg, M., et al. (2010). Interference with bile salt export pump function is a susceptibility factor for human liver injury in drug development. *Toxicol. Sci.* 118, 485–500. doi: 10.1093/toxsci/ikf269
- Nicolaou, M., Andress, E. J., Zolnericiks, J. K., Dixon, P. H., Williamson, C., and Linton, K. J. (2012). Canalicular ABC transporters and liver disease. *J. Pathol.* 226, 300–315. doi: 10.1002/path.3019
- Ozer, J., Ratner, M., Shaw, M., Bailey, W., and Schomaker, S. (2008). The current state of serum biomarkers of hepatotoxicity. *Toxicology* 245, 194–205. doi: 10.1016/j.tox.2007.11.021
- Padda, M. S., Sanchez, M., Akhtar, A. J., and Boyer, J. L. (2011). Drug-induced cholestasis. *Hepatology* 53, 1377–1387. doi: 10.1002/hep.24229
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* 12, 2825–2830.
- Pinto, M., Trauner, M., and Ecker, G. F. (2012). An *in silico* classification model for putative ABCC2 substrates. *Mol. Inform.* 31, 547–553. doi: 10.1002/minf.201200049
- Pires, D. E. V., Blundell, T. L., and Ascher, D. B. (2015). pkCSM: predicting small-molecule Pharmacokinetic And Toxicity Properties Using Graph-Based Signatures. *J. Med. Chem.* 58, 4066–4072. doi: 10.1021/acs.jmedchem.5b00104
- Sahigara, F., Ballabio, D., Todeschini, R., and Consonni, V. (2013). Defining a novel k-nearest neighbours approach to assess the applicability domain of a QSAR model for reliable predictions. *J. Cheminform.* 5:27. doi: 10.1186/1758-2946-5-27
- Stieger, B. (2011). “The role of the sodium-taurocholate cotransporting polypeptide (NTCP) and of the bile salt export pump (BSEP) in physiology and pathophysiology of bile formation,” in *Handbook of Experimental Pharmacology* (Berlin; Heidelberg: Springer), 205–259. doi: 10.1007/978-3-642-14541-4_5
- Szakács, G., Annereau, J.-P., Lababidi, S., Shankavaram, U., Arciello, A., Bussey, K. J., et al. (2004). Predicting drug sensitivity and resistance:

- profiling ABC transporter genes in cancer cells. *Cancer Cell* 6, 129–37. doi: 10.1016/j.ccr.2004.06.026
- U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (2012). *Guidance for Industry. Drug Interaction Studies Study Design, Data Analysis, Implications for Dosing, and Labeling Recommendations*. Rockville, MD.
- Warner, D. J., Chen, H., Cantin, L.-D., Kenna, J. G., Stahl, S., Walker, C. L., et al. (2012). Mitigating the inhibition of human bile salt export pump by drugs: opportunities provided by physicochemical property modulation, *in silico* modeling, and structural modification. *Drug Metab. Dispos.* 40, 2332–41. doi: 10.1124/dmd.112.047068

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Montanari, Knasmüller, Kohlbacher, Hillisch, Baierová, Grandits and Ecker. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.