Toward predictive models for drug-induced liver injury in humans: are we there yet?

Drug-induced liver injury (DILI) is a frequent cause for the termination of drug development programs and a leading reason of drug withdrawal from the marketplace. Unfortunately, the current preclinical testing strategies, including the regulatory-required animal toxicity studies or simple *in vitro* tests, are insufficiently powered to predict DILI in patients reliably. Notably, the limited predictive power of such testing strategies is mostly attributed to the complex nature of DILI, a poor understanding of its mechanism, a scarcity of human hepatotoxicity data and inadequate bioinformatics capabilities. With the advent of high-content screening assays, toxicogenomics and bioinformatics, multiple end points can be studied simultaneously to improve prediction of clinically relevant DILIs. This review focuses on the current state of efforts in developing predictive models from diverse data sources for potential use in detecting human hepatotoxicity, and also aims to provide perspectives on how to further improve DILI prediction.

**Minjun Chen[1],
Halil Bisgin[1],
Lillian Tong[1,2],
Huixiao Hong[1],
Hong Fang[3],
Jürgen Borlak[‡4]
& Weida Tong*[‡1]**

[1]Division of Bioinformatics & Biostatistics, National Center for Toxicological Research, The US Food & Drug Administration, Jefferson, AR, USA
[2]Division of Biology & Biological Engineering, California Institute of Technology, Pasadena, CA, USA
[3]Office of Scientific Coordination, National Center for Toxicological Research, The US Food & Drug Administration, Jefferson, AR, USA
[4]Center of Pharmacology & Toxicology, Hannover Medical School, Hannover, Germany
*Author for correspondence:
weida.tong@fda.hhs.gov
[‡]Authors contributed equally

Drug-induced liver injury (DILI) is a clinical event that can be associated with severe outcomes such as acute liver failure [1]. Approximately 1000 drugs, herbal products, vitamins and illicit compounds are associated with liver injury [2], albeit generally at a low frequency. These DILI events are mostly of an idiosyncratic nature and are host dependent [3,4]. DILI is also a major cause for the termination of drug development programs [5] and frequently results in regulatory actions including denied approval and black box warnings [6]. Notably, more than 50 approved drugs have been withdrawn from domestic markets in different countries worldwide owing to DILI [7,8].

Currently employed preclinical testing strategies do not predict DILI in patients reliably [9], with less than 55 and 25% of DILI drugs being predicted based on the regulatory animal toxicity studies and simple *in vitro* tests, respectively [10,11]. Major investments have therefore been made to reduce risk for DILI, including the search and validation of predictive biomarkers [12]. Diverse liver-related biological and chemical data have been deposited in the public domain to bridge the gap between DILI prediction and clinical outcome, enabling independent research teams to analyze the data. Turning data into knowledge remains a challenge, particularly with microarray-based toxicogenomic data sets. Likewise, very large relational databases can be generated by high-content screening (HCS) assays that simultaneously monitor multiple biological end points that become enormously valuable when combined with whole-genome toxicogenomics. However, these tests mostly rely on a limited number of drugs, which limits statistical power for developing robust predictive models. This review will focus on the current state of efforts to develop predictive models that can be used in preclinical studies for detecting DILI risk in humans.

Classification of drugs for DILI risk in humans

An improved understanding of DILI is vital for the development of accurate and useful predictive models [13]. DILI is alleged to be either a drug's inherent property or an idiosyncratic outcome peculiar to a case or individual; however, the value and validity of such a classification is questionable and no conclusive evidence supports that intrinsic and idiosyncratic DILIs arise by different modes of action [14,15]. Here, we focused on the classification of DILI risk in humans instead.

Annotating a drug's potential to cause DILI in humans is a huge undertaking in a clinical setting [16]. At least three attributes, including severity, causality and incidence, need to be considered when a drug's DILI risk in humans is assessed [8]. The difficulty of accurate annotations can be attributed to factors such as the diverse and complicated clinical manifestation of DILI, the lack of sensitive and specific biomarkers for DILI diagnosis, the difficulties in determining

causality and the serious under-reporting of DILI cases.

Some prominent data sources for the classification of human DILI risk are listed in TABLE 1, and each are either case report-based approaches, drug compendium-based approaches or a combination of these. Clinical investigators always prefer the case report-based approach. Suzuki *et al.* surveyed multiple well-documented and adjudicated DILI cases and developed a repository including 319 drugs causing overall liver injury, 107 drugs causing acute liver failure and 47 drugs suspended or withdrawn from the market [17]. Likewise, drug compendia such as the Physicians' Desk Reference are widely used as resources. Drug compendia are continually updated and provide judgmental information based on a body of comprehensive evidence and expert opinions. Chen *et al.* employed US FDA-approved drug labels to assess DILI risk in humans, categorizing drugs as of most-DILI-concern, less-DILI-concern or no-DILI-concern [8]. The DILI concern classification is in a straightforward manner based on the drug label section and severity of DILI. Despite imperfections and limitations, drug labels reflect the serious thoughts and consensus of numerous experts and constitute one major public data source that comprehensively integrates case reports and compendia [8].

Most published DILI classifications are of case reports and compendia. Xu *et al.* defined a drug to be DILI positive, if it was either withdrawn, not approved in the USA, issued with a boxed warning, marketed with warnings related to DILI or associated with a significant number (>10) of case reports of serious DILIs that met the criteria of the Hy's law, which estimates risk of fatal DILI [18]. Sakatis *et al.* annotated a drug as DILI when at least 50 reports of clinical hepatotoxicity or three reports of life-threatening events were recorded in a bibliographic database of drug-related liver injuries and the Physicians' Desk Reference [19].

In an effort for assessing the consistency of DILI classifications, four data sets were selected (the Suzuki *et al.* [17], Chen *et al.* [8], Xu *et al.* [18], and Sakatis *et al.* [19] data sets) as indicated in SUPPLEMENTARY TABLE 1 (see online at www.future-medicine.com/doi/full/10.2217/BMM.13.146). The drugs qualified for analyses were only those that were classified in two or more data sets as well as belonging to the group of FDA-approved drugs. A drug was defined as equivocal classification when there was disagreement between any two data sets. Among the qualified drugs, 181 drugs (74%) have consistent DILI classification, and 64 drugs were equivocal. The equivocal drugs were further analyzed to identify in which label sections their DILI risk were disclosed. Some 15 drugs were in 'Warning & Precautions' (19% of the 77 'Warning & Precautions' drugs), 39 were in 'Adverse Reactions' (63% of the 62 'Adverse Reactions' drugs), and ten were in 'No Match' (24% of the 41 'No Match' drugs). Notably, the DILI lexeme 'Adverse Reactions' is quite debatable and accounts over 60% of the total 64 equivocal drugs.

Predictive models from homogenous data

The primary impediment to improving DILI prediction models and screening classifiers is the poor understanding of DILI mechanisms and pathogenesis. Models so far fall into three categories as determined by type of data used: first, chemical structure based; second, *in vitro* assay-based; and third, toxicogenomics based. The prominent predictive models selected from the literature are summarized in TABLE 2. Animal DILI models are not included since most are drug-specific studies (for a comprehensive review see [20]).

■ Chemical structure-based *in silico* models

Once the model has been developed, the chemical structure-based *in silico* model has the advantage of rapidly screening an unlimited number of chemicals at minimal cost. The basic assumption is that similar chemical structures have similar properties and toxicity profiles [21]. Use of structure-based models to screen for DILI has been limited, but efforts have nonetheless progressed [22] using both knowledge-based expert systems and quantitative structure–activity relationship (QSAR)-based approaches [23].

Knowledge-based models

The knowledge-based or expert system models are usually developed with commercial software such as Derek for Windows (Lhasa Limited) from which structure alerts are derived. The structure alerts are chemical substructures or motifs the expert system statistically associates with the biological end point of interest. The validity of the structural alerts needs to be assessed by experts in order to minimize and balance false-positive and false-negative predictions. Groups of chemicals with similar toxicity profiles are identified in order to ascertain the shared structural motifs associated with toxicity.

Table 1. Published classifications of drugs for drug-induced liver injury risk in humans.

Study (year)	Hepatotoxicity classification	Supporting evidence	Drugs/chemicals annotated (n)	Ref.
Greene et al. (2010)	AH (animal hepatotoxicity observed, not tested in humans); HH (human hepatotoxicity observed); WE (weak evidence of human hepatotoxicity [<10 case reports] but generally considered safe); and NE (no evidence of hepatotoxicity in any species)	Not defined	139 AH drugs; 273 HH drugs; 62 WE drugs; and 152 NE drugs	[26]
Suzuki et al. (2010)	Drugs associated with overall liver injury; acute liver failure; or suspended or withdrawn owing to hepatotoxicity	Databases created by ongoing prospective studies to collect cases of DILI or acute liver failure, databases created using regulatory resources, published literature and other public domains	319 drugs causing overall liver injury; 107 drugs causing acute liver failure; and 47 suspended or withdrawn drugs	[17]
Chen et al. (2011)	Most-DILI-concern (drugs were suspended, withdrawn or issued with a black box warning owing to hepatotoxicity or received warnings and precaution with moderate or severe DILI concern); less-DILI-concern (drugs were issued with warnings and precaution with mild DILI concern or only recorded hepatotoxicity in adverse reactions); and no-DILI-concern (no DILI mentioned in the drug label)	US FDA-approved drug label	137 most-DILI-concern drugs; 85 less-DILI-concern drugs; and 65 no-DILI-concern drugs (total of 287 drugs)	[8]
Sakatis et al. (2012)	Hepatotoxic drugs (at least 50 reports of clinical hepatotoxicity or three reports of life-threatening); and nonhepatotoxic drugs (no reports of clinical hepatotoxicity or only one report of mild hepatotoxicity [transient mildly elevated enzymes])	The bibliographic database of drug-related liver injuries and Physicians' Desk Reference	113 hepatotoxic drugs; and 110 nonhepatotoxic drugs	[19]
Zimmerman (1999)	Drugs associated with hepatotoxicity	Literature	~500 drugs	[84]
Guo et al. (2005)	3 (clear literature evidence of life-threatening hepatotoxicity); 2 (multiple case reports or significant liver injuries); 1 (no significant liver damage has been reported); and 0 (no information about hepatotoxicity)	American Hospital Formulary Service®, US Pharmacopeial Convention, Facts & Comparisons®, Physicians' Desk Reference and Clinical Pharmacology	175 drugs	[85]
Xu et al. (2008)	DILI positives (e.g., withdrawn from the market mainly owing to hepatotoxicity; not marketed in the USA owing to hepatotoxicity; received black box warnings from the FDA owing to hepatotoxicity; marketed with hepatotoxicity warnings in their labels; others that had well-known associations to liver injury and had a significant number (>10) of independent clinical reports of serious hepatotoxicity that meet the criteria of Hy's Law; and a small number of Pfizer internal drug candidates whose development were ceased mainly owing to hepatotoxicity concerns); and DILI negatives (drugs that do not meet any of the above positive criteria)	Drug label, expert opinion and cases reported in the literature	200 DILI positives; and 144 DILI negatives	[18]
Hoofnagle et al. (2013)	Drugs associated with hepatotoxicity	Literature	~650 drugs, dietary supplements and herbal products	[79]

DILI: Drug-induced liver injury.

Table 2. Predictive models for assessing drug-induced liver injury risk in humans.

Data used	Classification algorithm	Performance assessment approach	Performance reported	Ref.
Knowledge-based models				
Structure alerts	Derek for Windows	244 drugs	Not reported	[22]
Structure alerts	Derek for Windows	137 drugs for EV	40% SEN	[25]
Structure alerts	Derek for Windows	623 compounds for EV	56% ACC; 46% SEN; and 73% SPEC	[26]
Structure alerts	Toxmatch structural similarity tool	Not reported	Not reported	[21]
QSAR-based models				
2D molecular descriptor	Ensemble recursive partitioning	382 drugs for CV	CV: 76% ACC; 76% SEN; and 75% SPEC	[27]
		54 drugs for EV	EV: 81% ACC; 70% SEN; and 90% SPEC	
Radial distribution function molecular descriptors	Linear discriminant analysis	74 drugs for CV	CV: 84% ACC; 78% SEN; and 90% SPEC	[28]
		13 drugs for EV	EV: 82% ACC	
Molecular descriptors	Four commercial QSAR programs	~1600 drugs for CV	CV: 39% SEN; and 87% SPEC	[29]
		18 drugs for EV	EV: 89% SEN	
MolConnZ and Dragon molecular descriptor	k-nearest neighbor	37 drugs for EV	84% ACC; 74% SEN; and 94% SPEC	[30]
2D fragments and Dragon molecular descriptors	Support vector machine	531 drugs for CV	CV: 62–68% ACCs	[31]
		18 compounds for EV	EV: 78% ACC	
ECFC_6 molecular descriptors	Linear discriminant analysis	295 compound for CV	CV: 59% ACC; 53% SEN; and 65% SPEC	[33]
		237 compounds for EV	EV: 60% ACC; 56% SEN; and 67% SPEC	
PaDEL molecular descriptor	Ensemble of mixed learning algorithms	1087 compounds for CV	CV: 68% ACC; 67% SEN; and 70% SPEC	[32]
		120 compounds for EV	EV: 75% ACC; 82% SEN; 65% SPEC	
ECFC_6 molecular descriptors	Bayesian models	Three data sets with 40–148 drugs for EV	EV: 60–70% ACCs	[34]
Mold2 chemical descriptor	Decision forest	197 drugs for CV	CV: 70% ACC	[35]
		Three data sets with 190–348 drugs for EV	EV: 62–69% ACCs	
In vitro *assay-based models*				
Five end points based on Hep G2 cell line	Logical OR algorithm	243 drugs for IV	80% SEN; and 90% SPEC	[41]
Four end points based on human primary hepatocyes	Decision tree	344 compounds for CV	50–60% SENs; and 95–100% SPECs	[18]
Four end points based on mouse liver mitochondria	Logical OR algorithm	114 compounds for IV	92% SEN; and 82% SPEC	[40]
Five end points based on Hep G2 cell line	Logical OR algorithm	78 drugs for IV	94% SEN; and 92% SPEC	[39]
Four end points based on human primary hepatocytes	Logical OR algorithm	45 drugs for IV	IV: 66% SEN; and 90% SPEC	[42]
		19 drugs for EV	EV: 100% SEN	

ACC: Accuracy; CV: Cross-validation; EV: External validation; GSH: Glutathione; IV: Interval validation; MDI: CYP450 metabolism-dependent inhibition; QSAR: Quantitative structure–activity relationship; SEN: Sensitivity; SPEC: Specificity.

Table 2. Predictive models for assessing drug-induced liver injury risk in humans (cont.).

Data used	Classification algorithm	Performance assessment approach	Performance reported	Ref.
In vitro *assay-based models* (cont.)				
Six end points based on Hep G2 cell line	Zone classifier	102 drugs for IV	45% SEN; and 100% SPEC	[43]
Covalent binding assay and daily dose	Ordinal logistic regression analysis	36 drugs for IV	80% ACC; 77% SEN; and 86% SPEC	[45]
GSH adduct formation, marked CYP450 MDI or covalent binding assay	Decision tree	223 drugs for IV	IV: 73% ACC; 45% SEN; and 90% SPEC	[19]
		10 compounds for EV	EV: 70% ACC	

ACC: Accuracy; CV: Cross-validation; EV: External validation; GSH: Glutathione; IV: Interval validation; MDI: CYP450 metabolism-dependent inhibition; QSAR: Quantitative structure–activity relationship; SEN: Sensitivity; SPEC: Specificity.

Robust structure alerts with mechanistic implications have been used as screening tools to identify potential toxins and to mitigate the risk of idiosyncratic drug toxicity [24].

Egan *et al.* reported 74 structure alerts for predicting liver toxicity using a collection of 244 molecules [22]. Most of the structure alerts were functional groups related to reactive or toxic metabolites. The authors did not evaluate the performance but stated that prediction accuracy of alerts was limited. Later, Marchant *et al.* derived structure alerts from both propriety and literature data [25]. Their alerts were able to identify 40% of 137 chemicals in one published data set, but the sensitivity dropped to 23% when challenged by another set of 300 hepatotoxins. The authors claimed that the performance could be further improved by using applicability domains defined by lipophilicity and molecular weight. Hewitt *et al.* utilized a published data set comprising of 951 compounds having annotated DILI risk to develop structural alerts for DILI prediction [21]. Mechanistic rationales for the alerts were proposed, but the predictive performance was not adequately assessed. Greene *et al.* developed 23 DILI structure alerts based on the literature knowledge and assessed its performance using 623 compounds from a Pfizer-developed data set [26]. An overall accuracy of 56% was reported with 46% sensitivity and 73% specificity.

QSAR-based models

QSAR models have been developed using supervised machine learning methods where chemical structure is encoded by calculated molecular descriptors. Given adequate chemicals both active and inactive, QSAR approaches can be faster since the need for expert opinion is averted.

Cheng and Dixon were among the first to report a QSAR model for DILI prediction [27]. Some 149 hepatotoxins with dose–response data and 233 nonhepatotoxins comprised the training set for a classifier developed with recursive partitioning trees. Leave-10%-out cross-validation yielded 76% accuracy and the external validation on a hold-out data set of 54 drugs was of similar order, that is, 81% accuracy. Cruz-Monteagudo *et al.* employed several different classifiers to develop a predictive model based on 74 selected drugs [28]. Internal cross-validation yielded accuracies of 78–86% for classifiers, and 82% accuracy for an external data set of 13 hepatotoxins. Matthews *et al.* developed models for five types of DILI using QSAR programs based on approximately 1600 chemical structures [29]. Their best model exhibited 39% sensitivity and 87% specificity estimated from internal cross-validation and 89% sensitivity on an external validation set of 18 hepatotoxic drugs. Roger *et al.* reported a QSAR model based on approximately 200 selected drugs from the FDA's human liver adverse event database [30]. The model yielded 84% accuracy, 74% sensitivity and 94% specificity for predicting composite liver enzyme score assessed on an external validation set of 37 drugs. Fourches *et al.* developed QSAR classification models with a training set of 248 hepatotoxins and 283 nonhepatotoxins, and reported cross-validation accuracies of 62–68%, and 78% accuracy for an external validation set with 18 compounds [31].

Despite the fact that some of the previous reports presented moderately high predictive accuracy, their external validation sets are often quite small, consisting of only 20–50 drugs. Among the published models with larger external validation sets, the predictive performance seems to be less favorable. For example, Liew

et al. reported a QSAR model for DILI prediction [32]. When applied to an external validation of 120 compounds, the model achieved 75% accuracy, 82% sensitivity and 65% specificity. Ekins *et al.* developed a predictive ligand-based Bayesian QSAR model for assessing human DILI using a training set of 295 compounds [33]. It achieved 60% accuracy with 56% sensitivity and 67% specificity on an external validation set with 237 drugs. Liu *et al.* developed a DILI prediction system based on a group of *in silico* predictive models using 13 selected DILI-relevant end points [34]. Accuracies of 60–70% were achieved when assessed by three external validation sets with a total of >200 drugs. Chen *et al.* developed a QSAR model using Mold2 molecular descriptors, a decision forest algorithm and a training set with 197 drugs [35]. The model achieved 69, 62 and 63% accuracies when assessed by three separate external validation sets having a total of approximately 500 drugs. Notably, the three studies (i.e., Liew *et al.* [32], Ekins *et al.* [33] and Chen *et al.* [35]) and except for Liu *et al.* [34] employed a relative balance of hepatotoxic-positive and -negative drugs in the validation sets.

Considering the applicability domain of a model, as determined by the chemical space of the training set, is a powerful way to ascribe confidence to predictions [36]. If a chemical's structure is within the applicability domain of the model, there will be greater confidence in a given prediction of an untested chemical. By contrast, if a chemical's structure is outside the domain, prediction confidence will be low. An applicability domain has been successfully applied to the prediction of end points with well-established mechanisms by attaching statistical confidence to individual predictions [37,38]; however, some studies reported no such benefit when used for DILI prediction [32,33,35]. For example, Liew *et al.* reported that chemical structure-based applicability domain did not delineate separate categories of high-confidence and low-confidence predictions, with corresponding higher and lower prediction accuracy, respectively [32], and these observations were confirmed independently [33,35]. Instead, Chen *et al.* proposed a therapeutic categories-based applicability domain approach and demonstrated its superiority in the prediction of DILI when assessed by three independent external validation sets [35].

■ *In vitro* assay-based models
Due to its requirement for a low quantity of drug substances and the possibility of high-throughput screening, *in vitro* assays have been widely used in preclinical testing. However, conventional cytotoxicity assays employing a single end point have poorly correlated (mostly <25% sensitivity) with human DILI outcomes [11], an unremarkable results given that the assay has little to do with putative mechanisms and pathogenesis of DILI. Simultaneous use of multiple *in vitro* models could improve results as a broader spectrum of molecular mechanisms would be encompassed, some of which might be associated with DILI. HCS assays that simultaneously measure multiple end points and associated cellular reactions offer an opportunity to detect changes of DILI-related molecular signals in live cells with higher sensitivity [39,40].

One of the first reports using HCS for hepatotoxicity employed the HepG2 (human hepatoma) cell line [41]. The HCS assay tested 243 drugs/chemicals that were a mixture of hepatotoxins, nontoxins and those known to be toxic in organs other than liver. Five mechanistic end points (i.e., intracellular calcium level, mitochondrial membrane potential [MMP], cell number, nuclear area and plasma membrane permeability) were measured. A predictive model based on a simple logistical combination of the measured end points reported 80% sensitivity and 90% specificity. Later, Xu *et al.* developed a primary human hepatocyte-based HCS assay and tested over 300 drugs/chemicals at a single concentration of 100-times human the therapeutic plasma concentration C_{max} [18]. Four related assays (nuclear and lipid content, intracellular glutathione, mitochondrial damage and oxidative stress) were measured after a 24-h cell incubation. A predictive model based on a random forest algorithm yielded sensitivities of 50–60% and specificities of 95–100%. Considering the rapid deterioration of primary hepatocytes cultured under conventional conditions, Khetani *et al.* developed a micropatterned co-culture liver model that enabled long-term dosing in functionally viable hepatocytes [42]. They tested 45 drugs along with four standard cytotoxic end points (i.e., glutathione, ATP level, albumin and urea secretion). The human primary hepatocytes model yielded 66% sensitivity and 90% specificity. More recently, Persson *et al.* reported a HepG2 assay based on six parameters (nuclei counts, nuclear area, plasma membrane integrity, lysosomal activity, MMP and mitochondrial area) [43]. They developed a zone classifier based on nuclei count, MMP and human C_{max}, and reported 45% sensitivity and 100% specificity when tested on 102 selected drugs.

Predictive models derived from other *in vitro* assays that improve prediction of DILI have also appeared. For example, covalent binding is a measure of chemically reactive metabolites, but alone is insufficient to discriminate positive from negative hepatotoxic compounds [24,44]. Nakayama *et al.* developed a zone classifier using an ordinal logistic regression analysis based on daily dose and covalent binding [45]. It correctly identified 17 out of 22 hepatotoxic drugs (77% sensitivity) and 12 out of 14 safe drugs (86% specificity). More recently, Sakatis *et al.* developed a testing strategy using a decision tree algorithm based on assays for covalent binding, daily dose, CYP450 metabolism inhibition and glutathione adducts [19]. The model was assessed using 113 hepatotoxic-positive and 110 -negative marketed drugs and reported 45% sensitivity and 90% specificity. Additionally, the *in vitro* models like those for screening intrahepatic cholestasis allow insight into the underlying mechanisms of DILI [46].

Notably, most of the *in vitro*-based predictive models required drug exposure data (i.e., human C_{max}). These data are commonly determined in human safety and efficacy studies and may possibly be predicted at the preclinical stage. Jones *et al.* predicted C_{max} with reasonable accuracy using *in silico* and *in vitro* data [47]. However, the predictions are far from perfect and tend to generally underestimate C_{max} [48].

■ Toxicogenomics-based models

The HCS assays alone can only offer limited sensitivities to identify liver liabilities of drugs mostly due to the mechanisms evolved in the available end points studied. In this regard, microarray-based technology enables unprecedented opportunities to comprehensively assess alterations in gene-expression profiles induced by hepatoxins. Consequently, the microarray technology was considered to be a groundbreaking addition to conventional toxicology assays at its inception [49,50]. However, the progress in developing toxicogenomic predictive models has not been as effective as envisioned, partly due to their high cost and the lack of advanced knowledge discovery and data mining tools [51]. Additionally, the transcriptomic approaches have been for end points not necessarily relevant to DILI, that is, the measured altered gene expression might or might not be associated with mechanisms involving hepatocellular injury. Two large-scale toxicogenomics data sets published recently, for example the Japanese toxicogenomics database [52] and

DrugMatrix [53], might enable real progress in developing DILI prediction models [51].

Several toxicogenomic studies have been performed for characterizing DILI; however, only very few models have been developed for predicting human DILI based on toxicogenomics data. Cha *et al.* developed a toxicogenomic model based on HepG2 cells to predict human DILI risk of NSAIDs [54]. The model identified 77 discriminatory genes derived from measurements for eight hepatotoxic and eight nonhepatotoxic drugs, and successfully classified four additional NSAIDs. Zhang *et al.* claimed that genomic biomarkers might not be more sensitive than the well-established markers of human DILI used in preclinical studies [55]. They demonstrated that the gene signatures derived from rats with abnormal ALT or bilirubin could predict human DILI, but the gene signatures derived from rats without ALT or bilirubin abnormal elevation cannot. Low *et al.* reported a toxicogenomic model derived from microarray data obtained from rat livers at 24 h after a single dose treatment, and the model was used to predict rat hepatotoxicity observed after 28 days of continued exposure [56]. The predictive model achieved a 76% correct classification rate (calculated as [sensitivity + specificity]/2) when assessed by fivefold cross-validation.

Metabolomics is yet another promising approach for discovering hepatotoxicity biomarkers. The approach has the major advantages of providing noninvasive sample collection and easy sample preparation and analysis [57,58]. A large-scale metabolomics study was reported that utilized the proton nuclear magnetic resonance spectra of rodent urine and serum to develop a predictive model for liver toxicity with 67% sensitivity and 77% specificity [59]. Circulating miRNAs also show potential to provide biomarkers of human DILI [60,61]. The miRNAs (miR-122 and miR-192) were observed to be significantly increased in the acetaminophen-induced liver injury patients compared with the healthy controls [62].

Predictive models from heterogeneous data

Most published models were developed from a homogenous data source, that is, chemical structure-based, *in vitro* assay-based or toxicogenomics-based data. Heterogeneous models employ one or more data types in an integrative approach. Rusyn *et al.* reviewed the efforts to predict *in vivo* toxicity by integrating chemical structure descriptors with *in vitro* assay

parameters to attain more predictive accuracy [63]. The heterogeneous models can be categorized into data integration or model integration modes in accordance with the objectives laid out.

■ Data integration

The data integration approach pools data from different sources prior to model development. Data integration from multiple sources can be challenging but provides opportunities to improve a model's predictive power. For example, Chen *et al.* analyzed 164 FDA-approved oral medications and showed an association of high daily doses (≥100 mg/day) and lipophilicity (partition coefficient [logP] ≥3) with significant risk for DILI, thus defining a 'rule-of-two' [64]. This principle was further applied and validated using an independent set of 179 oral medications, drug pairs with similar chemical structures and molecular targets, but different DILI potential, and to clinical case studies with complex comedication regimes. The authors demonstrated that the 'rule-of-two' caused fewer false positives than daily dose alone as a means of assessing risk for DILI in humans. Zhu *et al.* combined molecular descriptors from chemical structure and biological descriptors from *in vitro* assays to develop a hybrid model [65]. A fivefold cross-validation based on 156 DILI-positive and 136 DILI-negative compounds indicated that the hybrid model performed slightly better than the QSAR model alone. Moreover, Low *et al.* utilized chemical descriptors and toxicogenomic profiles to develop a hybrid model for 127 drugs to predict rat hepatotoxicity [56]. Assessed by a fivefold cross-validation, the hybrid model reported a correct classification of 68–77%, while the QSAR and toxicogemonic models alone reported a correct classification of 61 and 76%, respectively. The authors emphasized that the use of chemical and biological descriptors enriched mechanistic understanding and thus the model development.

Directly pooling data does not always work well for data integration, especially when the data structures of multiple data sources are different. Hence, an adaptor that does not compromise data integrity needs to be developed. Another type of indirect integration approach was developed by largely relying on advanced algorithms instead of a simple combination that incorporates different data sources together as a joint model. Consequently, it has the advantage of retaining the data structure of different data sources and mostly offers a better prediction [66].

Zitnik and Zupan developed a matrix factorization-based data integration approach to predict DILI potential of drugs from a gene-expression data set [101]. The model was derived from data of 29 different sources and achieved better accuracy than the conventional methods when assessed by cross-validation.

■ Model integration

Alternatively, models can be integrated after being independently developed from multiple data sources. The model integration approaches, including model consensus, have the capacity to improve predictive power [63,67]. Chen *et al.* reported a testing strategy by combining the 'rule-of-two' model with data from an *in vitro* HCS assay [Chen M, Tung C, Shi Q *et al.* Improve the prediction of *in vitro* assays for drug-induced liver injury by the integration of the 'rule-of-two'. (2013), Submitted]. Specifically, a primary rat hepatocyte-based HCS assay was used to further investigate the negative predictions from the 'rule-of-two' model. Assessed by a set of 73 drugs with clear evidence for clinically relevant DILI, the integrated model improved the accuracies by approximately 10% and reduced the number of drugs requiring *in vitro* test by approximately 20%, compared with the corresponding *in vitro* model alone.

Conclusion

Recent advances in preclinical testing strategies have improved our ability to identify drugs with risk for DILI. The construction of predictive models benefits from an integration of chemical structure, cellular end points, toxicogenomics data and data from multiple sources. These models can generate a high specificity between 90 and 95%, but perform less well for sensitivity (~50%). The limited power of DILI prediction is mostly attributed to the complex nature of DILI, a poor understanding of mechanisms, a scarcity of human hepatotoxicity data and insufficient bioinformatics capabilities. Altogether, the currently reported models have not demonstrated superiority of the predictive power over the regulatory-required animal toxicity studies, but the evolution of DILI study will continue to refine testing strategies.

Future perspective

The lack of an in-depth understanding of DILI mechanisms limits the development of improved DILI models. However, the continuing research efforts will result in knowledge gain that can be translated into novel testing strategies, thus

leading to better predictive models. For example, an *in vitro* approach that assays for the effects of cytokines was proposed to screen drugs with inflammation-associated idiosyncratic hepatotoxicity [68]. The assay was conceived from the observation that idiosyncratic hepatotoxicity for certain drugs has been reproduced in animals with mild inflammatory stress induced by bacterial LPS [9,20].

Despite the difficulty of performing mechanistic studies for idiosyncratic DILI, significant progress has been realized in identifying an association between carriers of specific HLA haplotypes and DILI events for some specific drugs [61]. For drugs such as flucloxacillin, abacavir, lapatinib, lumiracoxib, ticlopidine and ximelagatran that do not cause hypersensitivity, as seen for an immune-related response, an association between the HLA phenotype and risk for DILI was determined [12,69]. A large-scale genome-wide association study suggested a more significant role of common variants of immune-coding genes than those involved in drug metabolism or transport in the predisposition to DILI [70,71]. Similarly, the application of next-generation sequencing technologies (i.e., RNA-Seq) may help to identify rare functional variants underlying disease that are potentially associated with DILI events [71]. The clinical use of genetic association data in identifying DILI-susceptible patients is hurdled by the infrequent allele frequencies and a low incidence of drug-specific DILI events [71]. However, an understanding of the role of HLA in DILI events will stimulate the generation of more powerful predictive models for DILI [72].

Furthermore, enabling technologies based on pluripotent stem cells [73], organ-on-chips [74], zebrafish [75] and humanized mouse models [76] will provide powerful tools to investigate DILI mechanisms. Notably, the rapidly advancing research in stem cells raises possibility of using human pluripotent stem cells to provide more standardized hepatocytes for *in vitro* testing strategies [73]. The stem cell-derived hepatocytes represent a new cell source for the existing human primary hepatocyte model compared with thawed hepatocytes after cryopreservation and freshly-isolated hepatocytes. Until now, the stem cell-derived hepatotocyte systems are not fully validated and most of applications are still in research. Additionally, it is also promising to combine organotypic liver culture models based on hepatocytes and immune cells, the latter of which could be derived from patients who have experienced DILI [77].

The scarcity of human hepatotoxicity data in the public domain is another hindrance to developing predictive models. Due to the lack of sensitive and specific biomarkers for DILI diagnosis in clinical practice, DILI assessment is still largely a diagnosis of exclusion. The causality assessment of a DILI case is a laborious task and requires expert knowledge [78]. The need to comprehensively collect human hepatotoxicity data remains and the drug labeling-based approach seems to be more reliable, mostly owing to the intensive involvement of experts' knowledge and the reliance on comprehensive data collection [8]. Meanwhile, many efforts to collect the human hepatotoxicity data have been explored, including the NIH's LiverTox database [79,102] and the FDA's Liver Toxicity Knowledge Base (LTKB) [80,103].

Advanced computational techniques are also needed to mine the data and develop models with improved predictive power. Given the complexity of DILI mechanisms, a systematic approach is necessary to make full use of a wide range of information (experimental, clinical and computational) for the development of robust predictive models. Integrative models are promising, which can entrain specific algorithms to interrogate heterogeneous data and fuse them together in a joint model without distorting the modular structure of data from different sources. Network modeling is another promising approach to extract significant relationships among the features from heterogeneous data that benefit DILI mechanism studies [80]. Meanwhile, topic modeling has been demonstrated as a powerful text mining technique and demonstrated its utility in mining the FDA drug labels by extracting/discovering the relationships of different concepts (e.g., drugs, diseases and mechanisms) presented in the documents [81]. Furthermore, a unifying ontology to integrate histological and clinical observations for DILI lexeme was recently reported to improve the development of advanced knowledge discovery and data mining tools across different public repositories [82]. Overall, the advancement of bioinformatics tools will boost the development of DILI predictive models and help improve the understanding of its mechanisms.

For a complex end point such as DILI, a 'one-fits-all' model might fail to generate sufficient predictive performances. The success of applicability domains in the predictive model suggests that a proper stratification of drug space is essential. The therapeutic categories defined in the ontology of the WHO's

Anatomical Therapeutic Chemical Classification System were demonstrated to be a potential option [35]. Lastly, the predictive models should be validated, and their strengths and weaknesses should be fully assessed [83], which will be critical for their applications in drug development.

DILI studies are a concerted effort and involve multiple disciplines such as clinical hepatology, preclinical toxicology, molecular biology, and bioinformatics; an integration of different expert knowledge is the foundation for success. Several national and international programs have been initiated, including but not limited to: the AASLD-FDA-NIH-PhRMA – Hepatotoxicity Special Interest Group Conference [104], the DILI network project of the NIH [105], the DILIsym project of the Hamner Institutes for Health Sciences [106], the Virtual Liver Project (v-Liver™) of the US Environmental Protection Agency [107] and the Virtual Liver Network project funded by the German government [108]. All these efforts will improve our understanding of the complicated mechanisms of DILI to create better testing strategies that will reduce the time and cost of drug development and ultimately benefit patient safety.

Disclaimer

Acknowledgements

Financial & competing interests disclosure

Open Access

Executive summary

Background

- Insufficient predictive power of the current preclinical testing strategies has lead to a significant drug-induced liver injury (DILI) safety issue in the drug development program.

Classification of human DILI risk is vital for developing robust predictive models

- The consistencies among different DILI annotations are reasonably high.
- DILI lexeme 'Adverse Reactions' is debatable, which accounts for most of drugs with equivocal annotations.

Predictive models derived from homogenous or heterogeneous data

- The *in vitro*-based predictive models generally performed better than the chemical structure-based models.
- The integrative models from heterogeneous data sources demonstrated an improved predictive performance.

Novel predictive models on the horizon

- An association between HLA allele haplotypes and DILI risk for some drugs implicates the significant role of variants of immune-coding genes.
- Stem cell-derived hepatocytes represent a promising alternative model for DILI study.

Conclusion

- The currently reported models have not demonstrated superiority over the regulatory-required animal toxicity studies, but the evolution of DILI study will continue to refine testing strategies.

References

Papers of special note have been highlighted as:
- of interest
- of considerable interest

1   Ostapowicz G, Fontana RJ, Schiodt FV *et al.* Results of a prospective study of acute liver failure at 17 tertiary care centers in the United States. *Ann. Intern. Med.* 137(12), 947–954 (2002).

2   Reuben A, Koch DG, Lee WM. Drug-induced acute liver failure: results of a U.S. multicenter, prospective study. *Hepatology* 52(6), 2065–2076 (2010).

3   Zimmerman HJ. Drug-induced liver disease. *Clin. Liver Dis.* 4(1), 73–96, vi (2000).

4   Senior JR. What is idiosyncratic hepatotoxicity? What is it not? *Hepatology* 47(6), 1813–1815 (2008).

5   Ballet F. Hepatotoxicity in drug development: detection, significance and solutions. *J. Hepatol.* 26(Suppl. 2), 26–36 (1997).

6   Temple RJ, Himmel MH. Safety of newly approved drugs: implications for prescribing. *JAMA* 287(17), 2273–2275 (2002).

7   Fung M, Anna T, Kathy M, Wu J, Ken H, Edmundo M. Evaluation of the characteristics of safety withdrawal of prescription drugs from worldwide pharmaceutical markets – 1960 to 1999. *Drug Information Journal* 35(1), 293–317 (2001).

8   Chen M, Vijay V, Shi Q, Liu Z, Fang H, Tong W. FDA-approved drug labeling for the study of drug-induced liver injury. *Drug Discov. Today* 16(15–16), 697–703 (2011).

▪▪   **A novel drug-induced liver injury (DILI) annotation for drugs based on US FDA-approved drug labeling.**

9   Kaplowitz N. Idiosyncratic drug hepatotoxicity. *Nat. Rev. Drug Discov.* 4(6), 489–499 (2005).

10   Olson H, Betton G, Robinson D *et al.* Concordance of the toxicity of pharmaceuticals in humans and in animals. *Regul. Toxicol. Pharmacol.* 32(1), 56–67 (2000).

11   Xu JJ, Diaz D, O'Brien PJ. Applications of cytotoxicity assays and pre-lethal mechanistic assays for assessment of human hepatotoxicity potential. *Chem. Biol. Interact.* 150(1), 115–128 (2004).

12   Kaplowitz N, DeLeve LD. *Drug-Induced Liver Disease (Third Edition).* Academic Press, MA, USA (2013).

13   Przybylak KR, Cronin MT. *In silico* models for drug-induced liver injury – current status. *Expert Opin. Drug Metab. Toxicol.* 8(2), 201–217 (2012).

▪   **Good summary of the current *in silico* modeling for DILI.**

14   Corsini A, Ganey P, Ju C *et al.* Current challenges and controversies in drug-induced liver injury. *Drug Saf.* 35(12), 1099–1117 (2012).

▪▪   **Excellent summary of the current challenges and controversies in the study of DILI.**

15   Roth RA, Ganey PE. Intrinsic versus idiosyncratic drug-induced hepatotoxicity – two villains or one? *J. Pharmacol. Exp. Ther.* 332(3), 692–697 (2010).

16   Abboud G, Kaplowitz N. Drug-induced liver injury. *Drug Saf.* 30(4), 277–294 (2007).

17   Suzuki A, Andrade RJ, Bjornsson E *et al.* Drugs associated with hepatotoxicity and their reporting frequency of liver adverse events in VigiBase: unified list based on international collaborative work. *Drug Saf.* 33(6), 503–522 (2010).

18   Xu JJ, Henstock PV, Dunn MC, Smith AR, Chabot JR, de Graaf D. Cellular imaging predictions of clinical drug-induced liver injury. *Toxicol. Sci.* 105(1), 97–105 (2008).

▪   **Comprehensive evaluation of *in vitro*-based assays based on primary human hepatocytes for DILI prediction.**

19   Sakatis MZ, Reese MJ, Harrell AW *et al.* Preclinical Strategy to reduce clinical hepatotoxicity using *in vitro* bioactivation data for > 200 compounds. *Chem. Res. Toxicol.* 25(10), 2067–2082 (2012).

20   Roth RA, Ganey PE. Animal models of idiosyncratic drug-induced liver injury – current status. *Crit. Rev. Toxicol.* 41(9), 723–739 (2011).

21   Hewitt M, Enoch SJ, Madden JC, Przybylak KR, Cronin MTD. Hepatotoxicity: a scheme for generating chemical categories for read-across, structural alerts and insights into mechanism (s) of action. *Crit. Rev. Toxicol.* 43(7), 537–558 (2013).

22   Egan WJ, Zlokarnik G, Grootenhuis PDJ. *In silico* prediction of drug safety: despite progress there is abundant room for improvement. *Drug Discov. Today Technol.* 1(4), 381–387 (2004).

23   Greene N. Computer systems for the prediction of toxicity: an update. *Adv. Drug Deliv. Rev.* 54(3), 417–431 (2002).

24   Stepan AF, Walker DP, Bauman J *et al.* Structural alert/reactive metabolite concept as applied in medicinal chemistry to mitigate the risk of idiosyncratic drug toxicity: a perspective based on the critical examination of trends in the top 200 drugs marketed in the United States. *Chem. Res. Toxicol.* 24(9), 1345–1410 (2011).

25   Marchant CA, Fisk L, Note RR, Patel ML, Suarez D. An expert system approach to the assessment of hepatotoxic potential. *Chem. Biodivers.* 6(11), 2107–2114 (2009).

26   Greene N, Fisk L, Naven RT, Note RR, Patel ML, Pelletier DJ. Developing structure–activity relationships for the prediction of hepatotoxicity. *Chem. Res. Toxicol.* 23(7), 1215–1222 (2010).

▪   **Comprehensive evaluation of *in silico* models based on structure alerts for DILI prediction.**

27   Cheng A, Dixon SL. *In silico* models for the prediction of dose-dependent human hepatotoxicity. *J. Comput. Aided Mol. Des.* 17(12), 811–823 (2003).

28   Cruz-Monteagudo M, Cordeiro MN, Borges F. Computational chemistry approach for the early detection of drug-induced idiosyncratic liver toxicity. *J. Comput. Chem.* 29(4), 533–549 (2008).

29   Matthews EJ, Ursem CJ, Kruhlak NL *et al.* Identification of structure–activity relationships for adverse effects of pharmaceuticals in humans: part B. Use of (Q)SAR systems for early detection of drug-induced hepatobiliary and urinary tract toxicities. *Regul. Toxicol. Pharmacol.* 54(1), 23–42 (2009).

30   Rodgers AD, Zhu H, Fourches D, Rusyn I, Tropsha A. Modeling liver-related adverse effects of drugs using knearestneighbor quantitative structure–activity relationship method. *Chem. Res. Toxicol.* 23(4), 724–732 (2010).

31   Fourches D, Barnes JC, Day NC, Bradley P, Reed JZ, Tropsha A. Cheminformatics analysis of assertions mined from literature that describe drug-induced liver injury in different species. *Chem. Res. Toxicol.* 23(1), 171–183 (2010).

32   Liew CY, Lim YC, Yap CW. Mixed learning algorithms and features ensemble in hepatotoxicity prediction. *J. Comput. Aided Mol. Des.* 25(9), 855–871 (2011).

33   Ekins S, Williams AJ, Xu JJ. A predictive ligand-based Bayesian model for human drug induced liver injury. *Drug Metab. Dispos.* 38(12), 2302–2308 (2010).

▪   **Comprehensive evaluation of quantitative structure–activity relationship (QSAR) models for DILI prediction.**

34   Liu Z, Shi Q, Ding D, Kelly R, Fang H, Tong W. Translating clinical findings into knowledge in drug safety evaluation – drug induced liver injury prediction system (DILIps). *PLoS Comput. Biol.* 7(12), e1002310 (2011).

35   Chen M, Hong H, Fang H *et al.* Quantitative structure–activity relationship models for predicting drug-induced liver injury based on FDA-approved drug labeling annotation and using a large collection of drugs. *Toxicol. Sci.* 136(1), 242–249 (2013).

36   Tropsha A, Golbraikh A. Predictive QSAR modeling workflow, model applicability domains, and virtual screening. *Curr. Pharm. Des.* 13(34), 3494–3504 (2007).

37   Tong W, Hong H, Xie Q, Shi L, Fang H, Perkins R. Assessing QSAR limitations – a regulatory perspective. *Curr. Comput. Aided Drug Des.* 1(2), 195–205 (2005).

38   Tong W, Xie Q, Hong H *et al.* Using decision forest to classify prostate cancer samples on the basis of SELDI-TOF MS data: assessing chance correlation and prediction confidence. *Environ. Health Perspect.* 112(16), 1622–1627 (2004).

39   Tolosa L, Pinto S, Donato MT *et al.* Development of a multiparametric cell-based protocol to screen and classify the hepatotoxicity potential of drugs. *Toxicol. Sci.* 127(1), 187–198 (2012).

40   Porceddu M, Buron N, Roussel C, Labbe G, Fromenty B, Borgne-Sanchez A. Prediction of liver injury induced by chemicals in human with a multiparametric assay on isolated mouse liver mitochondria. *Toxicol. Sci.* 129(2), 332–345 (2012).

41   O'Brien PJ, Irwin W, Diaz D *et al.* High concordance of drug-induced human hepatotoxicity with in vitro cytotoxicity measured in a novel cell-based model using high content screening. *Arch. Toxicol.* 80(9), 580–604 (2006).

42   Khetani SR, Kanchagar C, Ukairo O *et al.* The use of micropatterned co-cultures to detect compounds that cause drug induced liver injury in humans. *Toxicol. Sci.* 132(1), 107–117 (2013).

43   Persson M, Løye AF, Mow T, Hornberg JJ. A high content screening assay to predict human drug-induced liver injury during drug discovery. *J. Pharmacol. Toxicol. Methods* 68(3), 302–313 (2013).

44   Usui T, Mise M, Hashizume T, Yabuki M, Komuro S. Evaluation of the potential for drug-induced liver injury based on *in vitro* covalent binding to human liver proteins. *Drug Metab. Dispos.* 37(12), 2383–2392 (2009).

45   Nakayama S, Atsumi R, Takakusa H *et al.* A zone classification system for risk assessment of idiosyncratic drug toxicity using daily dose and covalent binding. *Drug Metab. Dispos.* 37(9), 1970–1977 (2009).

46   Sakurai A, Kurata A, Onishi Y, Hirano H, Ishikawa T. Prediction of drug-induced intrahepatic cholestasis: *in vitro* screening and QSAR analysis ofdrugs inhibiting the human bile salt export pump. *Expert Opin. Drug Saf.* 6(1), 71–86 (2007).

47   Jones HM, Parrott N, Jorga K, Lave T. A novel strategy for physiologically based predictions of human pharmacokinetics. *Clin. Pharmacokinet.* 45(5), 511–542 (2006).

48   Poulin P, Jones RD, Jones HM *et al.* PHRMA CPCDC initiative on predictive models of human pharmacokinetics, part 5: Prediction of plasma concentration-time profiles in human by using the physiologically-based pharmacokinetic modeling approach. *J. Pharm. Sci.* 100(10), 4127–4157 (2011).

49   Afshari CA, Hamadeh HK, Bushel PR. The evolution of bioinformatics in toxicology: advancing toxicogenomics. *Toxicol. Sci.* 120(Suppl. 1), S225–S237 (2011).

50   Shi W, Bessarabova M, Dosymbekov D *et al.* Functional analysis of multiple genomic signatures demonstrates that classification algorithms choose phenotype-related genes. *Pharmacogenomics J.* 10(4), 310–323 (2010).

51   Chen M, Zhang M, Borlak J, Tong W. A decade of toxicogenomic research and its contribution to toxicological science. *Toxicol. Sci.* 130(2), 217–228 (2012).

52   Uehara T, Ono A, Maruyama T *et al.* The Japanese toxicogenomics project: application of toxicogenomics. *Mol. Nutr. Food Res.* 54(2), 218–227 (2010).

53   Ganter B, Snyder RD, Halbert DN, Lee MD. Toxicogenomics in drug discovery and development: mechanistic analysis of compound/class-dependent effects using the DrugMatrix database. *Pharmacogenomics* 7(7), 1025–1044 (2006).

54   Cha HJ, Ko M-J, Ahn S-M *et al.* Identification of classifier genes for hepatotoxicity prediction in non steroidalanti inflammatory drugs. *Mol. Cell. Toxicol.* 6(3), 247–253 (2010).

55   Zhang M, Chen M, Tong W. Is toxicogenomics a more reliable and sensitive biomarker than conventional indicators from rats to predict drug-induced liver injury in humans? *Chem. Res. Toxicol.* 25(1), 122–129 (2012).

56   Low Y, Uehara T, Minowa Y *et al.* Predicting drug-induced hepatotoxicity using QSAR and toxicogenomics approaches. *Chem. Res. Toxicol.* 24(8), 1251–1262 (2011).

57   Beger RD, Sun J, Schnackenberg LK. Metabolomics approaches for discovering biomarkers of drug-induced hepatotoxicity and nephrotoxicity. *Toxicol. Appl. Pharmacol.* 243(2), 154–166 (2010).

58   Schnackenberg LK, Chen M, Sun J *et al.* Evaluations of the *trans*-sulfuration pathway in multiple liver toxicity studies. *Toxicol. Appl. Pharmacol.* 235(1), 25–32 (2009).

59   Ebbels TMD, Keun HC, Beckonert OP *et al.* Prediction and classification of drug toxicity using probabilistic modeling of temporal metabolic data: the consortium on metabonomic toxicology screening approach. *J. Proteome Res.* 6(11), 4407–4422 (2007).

60   Wang K, Zhang S, Marzolf B *et al.* Circulating microRNAs, potential biomarkers for drug-induced liver injury. *Proc. Natl Acad. Sci. USA* 106(11), 4402–4407 (2009).

61   Grant LM, Rockey DC. Drug-induced liver injury. *Curr. Opin. Gastroenterol.* 28(3), 198–202 (2012).

62   Starkey Lewis PJ, Dear J, Platt V *et al.* Circulating microRNAs as potential markers of human drug-induced liver injury. *Hepatology* 54(5), 1767–1776 (2011).

63   Rusyn I, Sedykh A, Low Y, Guyton KZ, Tropsha A. Predictive modeling of chemical hazard by integrating numerical descriptors of chemical structures and short-term toxicity assay data. *Toxicol. Sci.* 127(1), 1–9 (2012).

64   Chen M, Borlak J, Tong W. High lipophilicity and high daily dose of oral medications are associated with significant risk for drug-induced liver injury. *Hepatology* 58(1), 388–396 (2013).

■■   A new rule for DILI prediction.

65   Zhu X, Sedykh A, Liu S. Hybrid *in silico* models for drug-induced liver injury using chemical descriptors and *in vitro* cell-imaging information. *J. Appl. Toxicol.* doi:10.1002/jat.2879 (2013) (Epub ahead of print).

66   Gevaert O, De Smet F, Timmerman D, Moreau Y, De Moor B. Predicting the prognosis of breast cancer by integrating clinical and microarray data with Bayesian networks. *Bioinformatics* 22(14), e184–e190 (2006).

67   Chen M, Shi L, Kelly R, Perkins R, Fang H, Tong W. Selecting a single model or combining multiple models for microarray-based classifier development? – A comparative analysis based on large and diverse datasets generated from the MAQC-II project. *BMC Bioinformatics* 12(Suppl. 10), S3 (2011).

68   Cosgrove BD, King BM, Hasan MA *et al.* Synergistic drug–cytokine induction of hepatocellular death as an *in vitro* approach for the study of inflammation-associated idiosyncratic drug hepatotoxicity. *Toxicol. Appl. Pharmacol.* 237(3), 317–330 (2009).

69   Liss G, Rattan S, Lewis JH. Predicting and preventing acute drug-induced liver injury: what's new in 2010? *Expert Opin. Drug Metab. Toxicol.* 6(9), 1047–1061 (2010).

70   Urban TJ, Shen Y, Stolz A *et al.* Limited contribution of common genetic variants to risk for liver injury due to a variety of drugs. *Pharmacogenet. Genomics* 22(11), 784 (2012).

71   Urban TJ, Goldstein DB, Watkins PB. Genetic basis of susceptibility to drug-induced liver injury: what have we learned and where do we go from here? *Pharmacogenomics* 13(7), 735–738 (2012).

72   Russmann S, Jetter A, Kullak-Ublick GA. Pharmacogenetics of drug-induced liver injury. *Hepatology* 52(2), 748–761 (2010).

73   Kia R, Sison RLC, Heslop J *et al.* Stem cell-derived hepatocytes as a predictive model for drug-induced liver injury: are we there yet? *Br. J. Clin. Pharmacol.* 75(4), 885–896 (2013).

74   Huh D, Hamilton GA, Ingber DE. From 3D cell culture to organs-on-chips. *Trends Cell Biol.* 21(12), 745–754 (2011).

75   Hill A, Mesens N, Steemans M, Xu JJ, Aleo MD. Comparisons between *in vitro* whole cell imaging and *in vivo* zebrafish-based approaches for identifying potential human hepatotoxicants earlier in pharmaceutical development. *Drug Metab. Rev.* 44(1), 127–140 (2012).

76 Dorner M, Horwitz JA, Robbins JB *et al.* A genetically humanized mouse model for hepatitis C virus infection. *Nature* 474(7350), 208–211 (2011).

77 Soldatow VY, LeCluyse EL, Griffith LG, Rusyn I. *In vitro* models for liver toxicity testing. *Toxicol. Res.* 2(1), 23–39 (2013).

78 Rockey DC, Seeff LB, Rochon J *et al.* Causality assessment in drug-induced liver injury using a structured expert opinion process: comparison to the Roussel-Uclaf causality assessment method. *Hepatology* 51(6), 2117–2126 (2010).

79 Hoofnagle JH, Serrano J, Knoben JE, Navarro VJ. LiverTox: a website on drug-induced liver injury. *Hepatology* 57(3), 873–874 (2013).

80 Chen M, Zhang J, Wang Y *et al.* Liver Toxicity Knowledge Base (LTKB) – a systems approach to a complex endpoint. *Clin. Pharmacol. Ther.* 95(5), 409–412 (2013).

81 Bisgin H, Liu Z, Fang H, Xu X, Tong W. Mining FDA drug labels using an unsupervised learning technique-topic modeling. *BMC Bioinformatics* 12(Suppl. 10), S11 (2011).

82 Wang Y, Lin Z, Liu Z *et al.* A unifying ontology to integrate histological and clinical observations for drug-induced liver injury. *Am. J. Pathol.* 182(4), 1180–1187 (2013).

83 Thomas CE, Will Y. The impact of assay technology as applied to safety assessment in reducing compound attrition in drug discovery. *Expert Opin. Drug Discov.* 7(2), 109–122 (2012).

84 Zimmerman HJ. *Hepatotoxicity: the Adverse Effects of Drugs and Other Chemicals on the Liver.* Lippincott, Williams & Wilkins, PA, USA (1999).

85 Guo JJ, Wigle PR, Lammers K, Vu O. Comparison of potentially hepatotoxic drugs among major US drug compendia. *Res. Social Adm. Pharm.* 1(3), 460–479 (2005).

■ Websites

101 Zitnik M, Zupan B. Matrix factorization-based data fusion for drug-induced liver injury prediction (2013). http://dokuwiki.bioinf.jku.at/lib/exe/fetch.php/zitnik.pdf

102 National Instituite of Diabetes and Digestive and Kidney Diseases and National Library of Medicine. LiverTox database. http://livertox.nih.gov (Accessed 25 October 2013)

103 The FDA's National Center for Toxicological Research. Liver Toxicity Knowledge Base. www.fda.gov/ScienceResearch/BioinformaticsTools/LiverToxicityKnowledgeBase/default.htm (Accessed 25 October 2013)

104 The AASLD-FDA-NIH-PhRMA – Hepatotoxicity Special Interest Group Conference. www.fda.gov/drugs/scienceresearch/researchareas/ucm071471.htm (Accessed 25 October 2013)

105 NIH. The DILI network project. www.fnih.org/work/programs-development/dilin (Accessed 25 October 2013)

106 DILI-sim Initiative – The Hamner Institutes for Health Sciences. www.thehamner.org/institutes-centers/institute-for-drug-safety-sciences/dili-sim-initiative (Accessed 25 October 2013)

107 US Environmental Protection Agency. The Virtual Liver project. www.epa.gov/ncct/virtual_liver (Accessed 25 October 2013)

108 The German government. The Virtual Liver Network project. www.virtual-liver.de/wordpress/en/the-virtual-liver/videos (Accessed 25 October 2013)