Quantitative Systems Toxicology Approaches to Understand and Predict Drug-Induced Liver Injury

Paul B. Watkins, MD

INTRODUCTION

Quantitative systems toxicology (QST) uses mathematical equations to recapitulate relevant pathways whereby drugs or other chemicals can cause death to cells and organs. The major QST effort currently is focused on drug-induced liver injury (DILI) is the DILI-sim Initiative. This is a public-private partnership established in 2011 to understand and predict the liver safety liability of new drug candidates. It involves

Disclosure: Dr P.B. Watkins chairs the Scientific Advisory Committee for the DILI-sim Initiative and receives compensation for this role. He also has a financial interest in the spinoff company DILIsym Services, Inc, which is a subsidiary of Simulations Plus.

Funding: No funding was received for this work.

Institute for Drug Safety Sciences, Eshelman School of Pharmacy, The University of North Carolina at Chapel Hill, 6 Davis Drive, PO Box 12137, Research Triangle Park, NC 27709, USA

E-mail address: pbwatkins@email.unc.edu

KEYWORDS

- DILIsym
- DILI
- QST
- Simulation
- Modeling

KEY POINTS

- The DILI-sim Initiative is a public-private partnership that has applied quantitative systems toxicology modeling to develop software (DILIsym®) that has improved mechanistic understanding of DILI.
- DILIsym incorporates pharmacokinetics and ability to alter key hepatocyte pathways to predict the frequency and severity of liver injury by drugs in simulated patient populations.
- Although DILIsym has been largely tested on drugs whose liver safety liability is already established, clinical trials are ongoing that will test its ability to prospectively predict liver safety before clinical trials are conducted.
- DILIsym also has been useful in optimizing interpretation of traditional liver chemistry tests and is incorporating new and promising biomarkers of liver injury.
- With further refinement of DILIsym, its predictions of liver safety may reduce the size of clinical trials required to establish liver safety and also may be useful in the clinic in managing DILI risk.
scientists from academia, the Food and Drug Administration (FDA), and pharmaceutical companies and has funding commitments until at least 2021. Partners in the Initiative vote to prioritize directions for the modeling, assuring that the software addresses the most pressing needs in drug development.

In the DILI-sim Initiative, the pathways whereby drugs can injure the liver are represented using differential equations in submodels, which are connected with the outcome of hepatocyte death and release of biomarkers into serum. Fig. 1 gives an overview of the submodels. DILIsym is the brand name of the evolving model, which is currently in version 8A. There are mouse, rat, and dog as well as human versions of the model.3,4 The first drug modeled by the initiative was acetaminophen where oxidative stress could account for toxicity observed with overdose in rodents and man. The modeling was used to propose the optimal protocol for treatment of acetaminophen overdoses with N-acetyl cysteine.5 The modeling was also then used to evaluate several hypotheses for why the isomer of acetaminophen, 3’-hydroxyacetanilide (AMAP), which also generates reactive metabolites, is much less toxic than acetaminophen in mice.6

The continued development of DILIsym has been based on data from many exemplar drugs with varying liver safety profiles, including drugs that had discordant results in preclinical and clinical testing. Exemplar drugs chosen for modeling included tenofovir, 12–15 ranitidine, 16–18 lamotrigine, 19–21 and other drugs that have undergone clinical and preclinical testing. These drugs represent a range of adverse drug reactions that occur with varying prevalence in preclinical and clinical testing, including intrinsic and extrinsic adverse drug reactions and idiosyncratic adverse drug reactions. The modeling suggests that at least for some of these adverse drug reactions, the toxicology of the drug can be modeled by different submodels. The continued modeling supports the use of DILIsym as a tool to predict hepatotoxicity. Further work is needed to test the mechanistic predictions of DILIsym in preclinical and clinical settings. The modeling has suggested that the 3 mechanisms outlined in thick boxes can account for hepatotoxicity in rats and man for more than 80% of the drugs in the validation cohort tested to date. The current version of the model also includes injury to cholangiocytes by inhibiting MDR3-mediated secretion of phospholipids into bile.
have included drugs where relevant data were publicly available and also unpublished data on drugs provided by the industry partners. To recapitulate the known safety profile of each exemplar drug, the model parameters were optimized. Once the model was optimized in this way, the Initiative began testing a new validation set of drugs where the preclinical and clinical safety profiles were known. As of May 2019, 68 molecules have been prospectively tested, with an 80% success in correctly identifying the presence or absence of a liver safety liability at the administered dosing (Brett Howell, personal communication, 2019). Among the 20% failures, all but 1 were predictions of safety with drugs that had exhibited some degree of hepatotoxicity (ie, false negatives).

DATA INPUTS TO DILIsym

The way DILIsym is typically used to assess the liver safety of a drug is illustrated in Fig. 2. Physiologically based pharmacokinetic (PBPK) modeling is created using available pharmacokinetic data and other relevant data (eg, liver-to-blood ratio of radioactivity in a rodent mass balance study) to estimate the time-dependent exposure of the drug outside and inside the hepatocyte. If the drug is a known substrate for uptake or efflux transporters, this fact is also taken into consideration in the PBPK model.

The properties of the drug relevant to hepatotoxicity are then assessed with in vitro systems. To screen for hepatocellular DILI potential, the drug is typically tested for its concentration-dependent ability to (1) inhibit bile acid transporters and thereby raise bile acid concentration in hepatocytes, (2) inhibit mitochondrial respiration, and (3) cause oxidative stress. If major metabolites are available, these typically also undergo these assays.

There are multiple hepatocyte transporters that can influence the intrahepatocyte concentration of bile acids,7 and the ability of a drug to inhibit each of these transporters is typically assayed. Some degree of inhibition of the bile salt export pump (BSEP) seems to be generally required to cause hepatotoxicity based on alterations in bile acid homeostasis, but the additional contribution resulting from inhibition of the basolateral efflux transporters MRP3 and MRP4 can be substantial. Conversely, inhibition of the major bile acid uptake pump, NTCP, would result in lowering of hepatocyte concentration of bile acids. Many drugs that inhibit efflux transporters also inhibit NTCP, creating a complex situation ideal for modeling.

The ability to inhibit mitochondrial respiration and to generate oxidative stress has been typically measured in a human hepatoma cell line, HepG2, using the Seahorse (Agilent Industries, Santa Clara, CA, United States) instrument and high content imaging, respectively. Because lipotoxicity is an infrequent mechanism of DILI, this property is assessed only if suspected. In addition to assessing the effect of the drug as a function of media concentration, the intracellular drug (and metabolite) concentration is also assessed using mass spectroscopy.

When modeling cholangiocyte injury, an assessment is made of the concentration-dependent ability of the drug and metabolites to inhibit a canalicular efflux transporter, multidrug-resistant protein 3 (MDR3). MDR3 transports into bile phospholipids that are incorporated into micelles. There are growing data to support the idea that reduction in biliary phospholipid reduces encapsulation of bile acids in micelles and that the resultant naked bile acids can be toxic to cholangiocytes.8 Cholangiocyte culture systems are currently being evaluated for the ability to generate relevant data, reflecting direct toxicity of the drug/metabolite to cholangiocytes.
The methods chosen to gather the data necessary for predictive modeling have been chosen by the DILI-sim Initiative partners because these methods are commercially available if not already up and running in their organizations.

**Data Outputs from DILIsym**

When the compound data collected are input into the model, together with estimates of the time-dependent concentration of the drug and major metabolites outside and inside hepatocytes, the model can then predict the time-dependent death of hepatocytes and hence the time-dependent release of biomarkers (typically ALT) into serum in an average patient. In addition, modeling can be conducted in Simpops® that have been created by changing parameters in the model to capture interpatient variation due to genetic or nongenetic factors. Thus, estimates can be made of the frequency as well as the extent of liver injury in a specific patient population targeted to receive the drug. If a drug causes elevations in liver injury biomarkers, hepatotoxicity can be minimized or eliminated in the Simpops by varying dose and liver chemistry monitoring parameters. This modeling has been helpful in designing clinical trials of new drug candidates (see text). (From Watkins PB. The DILI-sim Initiative: Insights into Hepatotoxicity Mechanisms and Biomarker Interpretation. Clinical and translational science. 2019;12(2):122-9; with permission.)
inside hepatocytes, the model then predicts the time-dependent death of hepatocytes and hence the time-dependent release of certain biomarkers into serum (see Fig. 1). The biomarker of most interest is generally serum alanine aminotransferase (ALT) because this is the most specific and sensitive among traditional biomarkers for hepatocyte death. Time-dependent changes in total bilirubin are also estimated based on the predicted loss of hepatocyte mass. In the model, bilirubin elevations to greater than 2-times the upper limit of normal (ULN) occur when the viable fraction of hepatocytes falls below 70%, a figure based on liver biopsy data obtained from patients experiencing liver injury due to acetaminophen overdose. Experimental biomarkers, including glutamate dehydrogenase, microRNA 122, full-length K18, and the caspase-cleaved fragment of cytokeratin 18, are also data outputs from the model, and incorporation of additional experimental biomarkers is planned.

The simplest output from DILIsym is predictions for an average healthy individual but DILIsym also can display predictions for simulated patient populations. This is done by varying parameters in the model to reflect interpatient variation in response resulting from genetic or nongenetic factors. The simulated populations (Simpops®) include patients with nonalcoholic steatohepatitis (NASH) and diabetes. Where possible, the model parameters for the patient-specific populations are varied based on experimental data, such as the reduction in activity in enzymes involved in mitochondrial oxidative phosphorylation observed in liver biopsies obtained from patients with NASH. Other parameters have been varied to fit data obtained from clinical trials involving the specific patient populations. New Simpops are planned for other patient populations, including children.

In addition to graphs for predicted liver chemistries over time, DILIsym can display predictions in evaluation of drug-induced serious hepatotoxicity (eDISH) format. eDISH is now a standard way the FDA evaluates liver safety of new drug candidates, generally from data obtained in phase 3 clinical trials. eDISH creates a graph, where, for each subject in a clinical trial, the observed peak serum ALT value is plotted along the X-axis and the observed peak serum bilirubin value along the Y-axis (ie, each subject is represented by a single point on the graph). Examples of eDISH graphs predicted by DILIsym are shown in Fig. 3. In this case, the modeling predicted that liver injuries, including severe liver injuries (ie, Hy’s law cases), would be encountered at high daily doses of the modeled drug. Safe dosing regimens, however, could be predicted.

Identifying Dominant Mechanisms Underlying Drug-Induced Liver Injury

Once liver safety liability of a drug is predicted by DILIsym, it is possible to identify which of the 3 mechanisms is contributing most to the predicted toxicity. This is done by simply turning off each mechanism in the model, 1 at a time, and observing the effect this has on the predicted frequency of serum ALT elevations in the Simpops. Typically, no 1 mechanism accounts for the predicted toxicity and there are instances where at least 2 mechanisms must be operative to produce any toxicity. There are as yet unpublished examples of where identifying the major mechanism underlying the toxicity of a drug has explained drug-drug interactions associated with increased frequency of elevations in serum ALT in clinical trials (Brett Howell, personal communication).

The prominence of the 3 mechanisms in accounting for toxicity is remarkable because none directly takes into account some DILI mechanisms that generally are recognized to be important, such as reactive metabolite production and endoplasmic reticulum stress. Such mechanisms may account for the approximately 20% failure rate of the current model predictions and addition of new mechanisms...
to DILIsym is likely in the future. It also is possible that there exist correlations with the mechanisms in the model such that those left out are indirectly taken into account. For example, a reactive metabolite may produce oxidative stress and oxidative stress can result in endoplasmic reticulum stress. Parent and major metabolites have been routinely tested in a human hepatoma cell line (Hep G2), which lack most of the drug metabolism capability of hepatocytes. The role of unrecognized metabolites, therefore, may account in part for DILIsym’s 20% prediction failure rate. The Initiative
has begun to collect mitochondrial inhibition and oxidative stress data in culture systems containing primary human hepatocytes or a hepatoma cell line that maintains most of the metabolic capacity of human hepatocytes (HepaRG) (with and without addition of nonparenchymal cells). In 1 case involving a molecule that was not predicted to be hepatotoxic by the current DILIsym inputs (see Fig. 2), time-dependent appearance of oxidative stress was noted in cultures of primary human hepatocyte, likely reflecting a role for unrecognized metabolites (Merrie Mosedale, personal communication).

**Mechanistic Insights**

Several major insights regarding hepatotoxicity mechanisms have evolved from the Initiative. These include importance of the mechanism underlying inhibition of bile acid transporters. When a drug is a competitive inhibitor of bile acid transporters, such as BSEP, as the hepatocyte concentration of the bile acids rises, the bile acids can out-compete the inhibitor to reduce the likelihood that toxic concentrations of bile acids in the hepatocyte will be achieved. In contrast, bile acids will be less able to override noncompetitive inhibition, and toxic concentrations are more likely to be achieved. The typically assessed concentration causing 50% inhibition of bile acid transport (IC50) does not identify mechanism of inhibition.

An example of where mechanism of inhibition of transporters was shown to be important in toxicity prediction involved an Amgen development program (Thousand Oaks, California). AMG 009 was a new drug candidate that demonstrated no liver safety signals in rat, dog, or monkey. Dose-dependent elevations in serum ALT, however, were observed in the phase 1 multiple dose-escalation clinical trial stopping clinical development. The program’s backup drug, AMG 853, which also had clean preclinical toxicology studies, was advanced into clinical development and no liver safety signals were observed. Retrospective investigation revealed that both drugs were potent inhibitors of BSEP, which is the major transporter of bile acids into bile. When assessed by IC50, AMG 853 was a more potent inhibitor of BSEP than AMG 009. AMG 853 was determined, however, to be a competitive inhibitor of BSEP whereas AMG 009 was a noncompetitive inhibitor of BSEP. There was no indication that either drug caused oxidative stress or interfered with mitochondrial function. The PBPK predictions of hepatocyte drug concentration together with the mechanistic inhibition for BSEP, NTCP, and other efflux transporters of bile acids were determined and loaded into DILIsym and the outputs examined. For AMG 009, the predicted time-dependent and dose-dependent elevations in serum ALT were similar to those observed in the phase 1 trial (ie, without any manipulation to the model parameters) (Fig. 4). No ALT elevations greater than 3-times the ULN were predicted for AMG 853 (not shown), consistent with the clinical trial experience, and this was demonstrated to be largely due the fact that AMG 853 was a competitive inhibitor of BSEP. In addition, DILIsym modeling of both compounds in the rat version of DILIsym predicted no liver injury consistent with the preclinical findings. The difference between rat and human toxicity of AMG 009 could be largely related to the inherently less hepatotoxic profile of bile acids in the rat. This was the first example of importance of the mechanism of transporter inhibition in assessing the hepatotoxic potential of a new drug candidate, but there have since been others.13

The modeling also has provided novel insight into mechanisms underlying species differences in susceptibility to hepatotoxicity18 and how structurally similar drug can have markedly different mechanisms of hepatotoxicity.17 A more complete review of mechanistic insights that have evolved from the DILI-sim Initiative is available.18
The Prospective Use of DILIsym in Drug Development

When the model predicts a new drug candidate will cause serum ALT elevations, it is possible to vary the dosage to predict regimens that should minimize or eliminate ALT elevations. In addition, if elimination of ALT elevations is not achieved at doses predicted to be effective, the model can predict the frequency of liver chemistry monitoring and appropriate stopping criteria based on ALT value to avoid serious liver injury. The use of DILIsym in this way has been applied to several new drugs, including an antibiotic.19

An increasingly frequent application of DILIsym has been the assessment of the safety of next-in-class drugs when first-in-class drugs have had liver safety issues. This has been demonstrated retrospectively with drug pairs known to have discordant liver safety profiles. For example, DILIsym predicted the Parkinson treatment tolcapone to be hepatotoxic and that the next-in-class drug entacapone was safe.20 Likewise, the model predicted troglitazone to be hepatotoxic and that the next-in-class drug pioglitazone was not.21 The discordant liver safety of these drug pairs was already established; the use of the term, prediction, refers to the fact the model results were not fitted to the known safety profiles but were the results that were produced by the model without manipulation. A crucial test of the model is predictions made for new drug candidates before they enter clinical trials.22 One such example involves 2 drugs to treat autosomal dominant polycystic kidney disease (ADPKD). The first-in-class drug, tolvaptan, was predicted by DILIsym to be potentially hepatotoxic in these patients whereas the next-in-class drug lixivaptan has been predicted by the model to be safe for the liver in the dosing proposed for this population.23 Clinical trials of lixivaptan in ADPKD patients are under way. There currently are several other prospective clinical trials of next-in-class drugs, with the dosing regimens predicted by DILIsym to be safe.

Application of DILIsym to Improve Biomarker Interpretation

As discussed previously, DILIsym predicts the time-dependent death of hepatocytes and the subsequent release into circulation of biomarkers, typically ALT, and also the rise in serum bilirubin as an indicator of global liver dysfunction (see Fig. 1). Alternatively, if serial assessments of serum ALT are available in an actual patients
experiencing hepatotoxicity, it is possible to fit the ALT kinetics to the model and, thereby, estimate the percent hepatocyte loss that the patient experienced. DILlsym actually creates a range of hepatocyte loss for a given ALT area under the curve, reflecting published variation in the content of ALT per hepatocyte and variation in the serum half-life of ALT. DILlsym was first used in this way to help interpret the significance of high peak serum ALT elevations (to 1000 IU/L) observed in healthy adults given a single injection of a toll-like receptor 5 agonist (a potential treatment of radiation sickness). Although the peak serum ALT elevations observed exceeded 20-times the ULN, the serum ALT values peaked quickly and then fell at approximately the published half-life of the enzyme. This suggested a short duration of hepatocyte death, and this was confirmed by the DILlsym modeling. The peak hepatocyte fraction lost was estimated to be less than 5% in the most affected volunteer. In more prolonged liver injuries, it is important that regeneration of hepatocytes in response to hepatocyte loss is built into DILlsym, such that the functioning hepatocyte mass is predicted at any point during and after resolution of ongoing hepatocyte death. The modeling has shown that regeneration rate can be important in determining liver function, particularly during prolonged liver injuries.

An additional consideration is that the release of ALT per hepatocyte into blood is reduced during apoptosis versus necrosis and this is built into the model. By simply looking at the effect of switching mode of cell death in the model, it is possible to estimate the effect of the 2 cell death pathways on estimated hepatocyte loss. It also has been proposed that during a DILI event, the proportions of apoptosis versus necrosis can be estimated as the ratio of the caspase-cleaved fragment of cytokeratin 18 to the full-length K18, termed, apoptotic index, and this ratio is incorporated into DILlsym. The apoptotic index recently was used to estimate the peak loss of hepatocyte mass in healthy volunteers treated with alfa cimaglermin alfa,25 a biologic agent proposed to treat congestive heart failure. In this case, the apoptotic index measured in archived serum samples supported apoptosis as the primary mode of liver cell death. The patient with the highest serum ALT peak value was predicted by DILlsym to have a peak reduction in hepatocyte mass of 12.5%. This was an important observation because this subject also experienced a rise in serum bilirubin exceeding 2-times the ULN (ie, a Hy’s law case), prompting a clinical hold on the development program. According to DILlsym, the maximal 12.5% loss of hepatocytes was not sufficient to account for a rise in serum bilirubin. It was later shown that the rise in serum bilirubin may be explained by on-target effects of the drug independent of toxicity.26 This modeling was presented as part of the regulatory communications aimed at removing the clinical hold on further development of this drug.

In addition to loss-of-function hepatocyte mass, DILlsym can predict elevations in direct and/or indirect bilirubin due to drug/metabolite inhibition of bilirubin transporters or inhibition of UGT1A1.27 A more complete review of the use of DILlsym in biomarker interpretation is available.24

**Future Applications of DILlsym**

DILlsym has shown success in predicting dose-dependent hepatotoxicity, including species differences in susceptibility. A greater challenge is predicting the idiosyncratic DILI results now believed to frequently involve immune an attack on the liver.28 Kupffer cell and recruited macrophage activation (the innate immune response in [Fig. 1](#)) are built into DILlsym.29 Activation of innate immune responses in DILlsym not only promotes hepatocyte injury but also affects hepatocyte regeneration rates. It is now clear, however, that many cases of DILI result from an adaptive immune attack on the liver. Incorporation of adaptive immune responses in DILlsym has begun and is an area of
emphasis going forward. This modeling should provide novel insights into mechanisms underlying idiosyncrasy as well as hepatotoxicity observed with immune modulators, such as the checkpoint inhibitors.30

Although adaptive immune responses are clearly important, the current versions of DILIsym successfully predicted the liver safety liability of 3 drugs that cause delayed idiosyncratic DILI, tolvaptan,31 troglitazone,21 and TAK-875.13 This success supports the generally accepted concept that initiation of an adaptive immune attack requires drug-induced hepatocyte stress28 and that this stress may in at least some cases be caused by the mechanisms already incorporated into DILIsym. Moreover, DILIsym predicted the several-month latency to peak serum ALT values observed in the clinical trials of troglitazone.21 This may suggest that delayed presentation of idiosyncratic hepatotoxicity can occur without involving an adaptive immune response. Regardless of the role of adaptive immunity in DILI, according to current concepts, drug regimens predicted by DILIsym to not cause hepatotoxicity should have a reduced chance of causing idiosyncratic DILI.

Looking farther into the future, DILIsym modeling may be able to reduce the size of clinical trials needed to establish liver safety. Just like modeling of drug interactions based on ability of drugs to induce or inhibit drug-metabolizing enzymes is increasingly accepted in place of performing drug-drug interaction clinical trials, DILIsym modeling in simulated patient populations ultimately may be accepted by regulators in place of actual clinical safety trials. Finally, DILIsym may someday be useful to clinicians in managing liver safety risks in their patients. If DILIsym modeling has already been performed for a specific drug, it may be possible for a physician to access the model through a smartphone, input patient specific data such as underlying diseases (eg, NASH) or concomitant medications, and in return be given a quantitative assessment of risk of DILI at various dosing regimens for that patient. DILIsym also may be useful in identifying the culprit drug in a patient with DILI receiving multiple drugs that have been modeled in DILIsym.

SUMMARY

DILIsym has evolved from a successful ongoing public-private partnership. It has provided novel insight into mechanisms underlying DILI, including interpatient variation in susceptibility to DILI. It is increasingly used in decision making within industry, and DILIsym modeling results are increasingly included in regulatory communications. It seems likely that DILIsym and QST efforts focused on other organs will improve the safety of new drugs while improving the efficiency of drug development.

REFERENCES


