Title

Evaluation of the relevance of DILI predictive hypotheses in early drug development: review of in vitro methodologies vs. BDDCS classification

Journal

Toxicology Research, 7(3)

ISSN

2045-452X

Authors

Chan, Rosa
Benet, Leslie Z

Publication Date

2018-05-08

DOI

10.1039/c8tx00016f

Peer reviewed

Evaluation of the Relevance of DILI Predictive Hypotheses in Early Drug Development: Review of *In Vitro* Methodologies *vs* BDDCS Classification

Rosa Chan[1] and Leslie Z. Benet[1]
rosa.chan@ucsf.edu
leslie.benet@ucsf.edu
[1]Department of Bioengineering and Therapeutic Sciences
Schools of Pharmacy and Medicine
University of California, San Francisco

To whom correspondence should be addressed:
Leslie Z. Benet
Department of Bioengineering and Therapeutic Sciences
Schools of Pharmacy and Medicine
University of California, San Francisco
533 Parnassus Avenue. Room U-68.
San Francisco, California 94143-0912
Phone: 1-415-476-3853
Fax: 1-415-476-8887
E-mail: leslie.benet@ucsf.edu

Keywords: Drug-induced liver injury, BDDCS, BSEP, daily dose, FDA hepatic liability

ABSTRACT

Drug-induced liver injury (DILI) is a major safety concern; it occurs frequently; it is idiosyncratic; it cannot be adequately predicted; and a multitude of underlying mechanisms has been postulated. A number of experimental approaches to predict human DILI have been proposed utilizing *in vitro* screening such as inhibition of mitochondrial function, hepatobiliary transporter inhibition, reactive metabolite formation with and without covalent binding, and cellular health, but they have achieved only minimal success. Several studies have shown total administered dose alone or in combination with drug lipophilicity to be correlated with a higher risk of DILI. However, it would be best to have a predictive DILI methodology early in drug development, long before the clinical dose is known. Here we discuss the extent to which Biopharmaceutics Drug Disposition Classification System (BDDCS) defining characteristics, independent of knowing actual drug pharmacokinetics/pharmacodynamics and dose, can be used to evaluate prior published predictive proposals. Our results show that BDDCS Class 2 drugs exhibit the highest DILI severity, and that all of the short-lived published methodologies evaluated here, except when daily dose is known, do not yield markedly better predictions than BDDCS. The assertion that extensively metabolized compounds are at higher risk of developing DILI is confirmed, but can be enhanced by differentiating BDDCS Class 2 from Class 1 drugs. Conclusion: Our published analyses suggest that comparison of proposed DILI prediction methodologies with BDDCS classification is a useful tool to evaluate the potential reliability of newly proposed algorithms, although BDDCS classification itself is not sufficiently predictive. Almost all of the predictive DILI metrics do no better than just avoiding BDDCS Class 2 drugs, although some early data with microliver platforms enabling long-enduring metabolic competency show promising results.

INTRODUCTION

Drug-induced liver injury (DILI) remains a major safety concern "due to its frequency of occurrence, idiosyncratic nature, poor prognosis, and multiple underlying mechanisms. Numerous experimental approaches have been published to improve human DILI prediction with modest success", as noted by Shah *et al.*[1] Idiosyncratic DILI (IDILI) is very complex. Most IDILI appears to be immune mediated, and reactive metabolites appear to be involved in most, but not all IDILI. Reactive metabolites are widely accepted as playing a pivotal role in the pathogenesis of idiosyncratic adverse drug reactions. While there are today well-established strategies for the risk assessment of stable metabolites within the pharmaceutical industry, there is still no consensus on reactive metabolite risk assessment strategies.[2] This is due to the complexity of the mechanisms of these toxicities as well as the difficulty in identifying and quantifying short-lived reactive intermediates such as reactive metabolites. In addition, there are probably several mechanisms by which a drug or reactive metabolite can induce an immune response. However, one might anticipate that new long-lived hepatocytes (i.e., liver-on-a-chip) might be better able to evaluate these conditions.

Across the pharmaceutical industry, systems of screening drug candidates have emerged that include transcriptomic profiling of animals in addition to animal pathology, assessment of covalent binding and glutathione (GSH) adducts in microsomal test systems and *in vivo*, inhibition of bile salt export pump (BSEP) *in vitro*, impairment of function of isolated animal mitochondria, and cell stress responses and viability in human hepatoma and hepatocyte culture systems. Several common themes emerge in all these test systems especially involving oxidative stress, mitochondrial impairment, covalent binding, and endoplasmic reticulum stress. It has

been proposed that these test systems have moderately strong predictive value for IDILI,[3–5] which we evaluate here. Others have examined combinations of mechanistic assays to better predict hepatotoxicity potential,[2,6] also evaluated here. Several studies have shown a correlation of total administered dose alone[7] or in combination with drug lipophilicity[8] with higher risk of DILI. However, it would be best to have a predictive DILI methodology early in drug development, long before the clinical dose is known.

Since liver injury has been reported with a large number of drugs, efforts have been undertaken to compile human hepatotoxicity data, including the National Institute of Health LiverTox Database[9] and the FDA Liver Toxicity Knowledge Base (LTKB).[10] These publicly available datasets have enabled development of new structure activity relationships for hepatotoxicity endpoints or triggered the development of knowledge-based and quantitative structure activity relationship (QSAR) models.[1,11,12]

We have reviewed the applicability of the Biopharmaceutics Drug Disposition Classification System (BDDCS) compared with presently proposed predictive procedures in evaluating DILI toxicity. Since its inception, the BDDCS has been useful in drug discovery for predicting routes of elimination, oral drug disposition, food effects on drug absorption, transporter effects on drug absorption, and potentially clinically significant drug interactions that may arise in the intestine, liver and brain.[13] We have shown that the BDDCS can be useful in predicting the potential for antiepileptic drugs to cause cutaneous adverse reactions[14] and in predicting DILI.[15] Most recently, we have shown that *in vitro* measures of BSEP inhibition, alone and together with other efflux transporters, provides no better prediction of DILI than just avoiding BDDCS class 2 drugs.[16] BDDCS's strong relationship between dose, metabolic

susceptibility, solubility and idiosyncratic DILI highlights the potential benefits of BDDCS as a comparison matrix for DILI prediction.

The BDDCS was developed in 2005 after Wu and Benet recognized that highly permeable compounds, as outlined by the Biopharmaceutics Classification System (BCS), were extensively metabolized, while poorly permeable drugs were primarily eliminated unchanged in the urine or bile.[17] BDDCS demonstrated that simple passive membrane permeability measures were highly selective in differentiating extensively vs. poorly metabolized drugs in humans. Drugs in the BDDCS are classified according to the membrane permeability rate and aqueous solubility. These characteristics have helped BDDCS define whether metabolic enzymes and/or transporters are clinically important. BDDCS features are demarcated by high and low values, classifying drugs into four categories. These classes are each associated with specific predictions regarding route of elimination and which interactions may be a clinical concern.[13]

Here we provide a review on the extent to which BDDCS defining characteristics, independent of knowing actual drug pharmacokinetics/pharmacodynamics and dose, can be used as a comparison baseline matrix of potential DILI adverse events with prior published predictive proposals.[13,15,16] We review the clinical impact of BDDCS in evaluating the severity of DILI warnings in drug labels approved by the Food and Drug Administration (FDA),[18] the withdrawal status due to adverse drug reactions (ADRs), the role of BSEP inhibition, maximum daily dosages prescribed, and *in vitro* toxicology assays applied to cover various mechanisms and toxicity endpoints associated with human DILI.[15]

Assessment of the BDDCS Classification on FDA Drug Labels Associated with DILI Hepatic Liability

In our previous work, we reported the BDDCS class relationship of hepatotoxicity between the different ADR categories by calculating the proportion of drugs in each FDA hepatic liability category, and each DILI severity category.[15] As depicted in Figure 1A, we observe that as the hepatic warning severity increases, the proportion of BDDCS Class 2 drugs increases and the proportions of both BDDCS Class 1 and 3 drugs decrease, all with highly significant trends. The "No mention" category is significantly different from all other categories, except for "Adverse Reactions." BDDCS Class 2 drugs exhibited the highest proportions in the following drug label sections: "Warning and Precautions" (45.6%, 36/79), "Boxed Warning" (47.2%, 17/36), "Withdrawn" (62.5%, 25/40) and "Discontinued" (83.3%, 5/6). Obviously, the number of drugs designated as exhibiting severe DILI increases as the ADR severity increases. That is, 15.9% (7/44) in the "Adverse Reactions" category, 36.7% (29/79) in the "Warning and Precautions" and 81.6% (31/38) of the drugs in the "Black Box Warning" are assessed to exhibit severe DILI.[15] In Figures 1B and 1C, the two BDDCS determinants (extent of metabolism and solubility) are examined. The percentages of poorly metabolized (Figure 1B) and of highly soluble (Figure 1C) drugs show statistically significant decreases with hepatic liability, while low solubility drugs increase significantly (Figure 1C) with hepatic liability. The percent of extensively metabolized drugs also increases with hepatic liability, but since almost 2/3 of "No mention" drugs are metabolized, it is apparent that extent of metabolism itself is not a discriminating parameter. Although greater extent of metabolism has been reported to significantly increase the potential of a compound to cause DILI,[19] this property alone is not able to distinguish compounds that are "No mention" of hepatic liability from those compounds exhibiting hepatic liability (See Figure 1B).

Our examination of the relationship between the BDDCS's determinant properties: solubility and extent of metabolism, led to some novel observations. Drugs belonging to BDDCS Class 1 and 3 exhibited a lower proportion of DILI severity. Drugs that are extensively metabolized and have low aqueous solubility, i.e., BDDCS Class 2 drugs, have the highest rates of DILI risk. BDDCS Class 2 drugs exhibited the highest proportions among the "Warning and Precautions", "Black Box Warning", "Withdrawn" and "Discontinued" categories. These are notably considered the most serious DILI risk categories (See Figure 1A). These findings demonstrate the importance of intrinsic drug properties as a potential factor for the development of a DILI event.

Drugs belonging to BDDCS Class 3 and 4 exhibited much lower proportions in the FDA hepatic liability (See Figure 1A). Moreover, BDDCS Class 3 and 4 drugs show a decreased risk of DILI related pathological events, such as liver aminotransferases increase and hyperbilirubinemia. However, we note the underrepresentation of BDDCS Class 4 drugs in the overall scheme of marketed approved drugs. Compounds with poor hepatic metabolism had been previously noted to be significantly less likely to cause hepatotoxicity.[19] Although a lack of hepatic metabolism does not assure total lack of hepatotoxicity, it indeed appears that BDDCS Class 3 and 4 drugs lead to a lower DILI severity.

Barton and co-workers[20] have previously discussed a new paradigm for navigating compound properties related to drug attrition. Optimizing the exposure of potent compounds at the desired site of action and in tissues associated with toxicity is fundamental to addressing attrition via efficacy and safety. Traditional oral drug space is well defined with respect to physicochemical properties and absorption, distribution, metabolism, excretion and toxicity (ADMET) risks but increased focus on ligand-lipophilicity efficiency, maximizing enthalpy

contributions and new target classes challenge this paradigm. Barton *et al.*[20] propose that BDDCS Class 3 compounds should be significantly more associated with drug attrition because they tend to be transporter substrates or inhibitors. Furthermore, they suggest that compounds that are substrates for transporters as being a toxicity liability. We completely disagree with this suggestion based on our analysis of DILI potential[15] and antiepileptic drugs' cutaneous adverse events.[14] Our analysis suggests that BDDCS Class 3 compounds exhibit less toxicity potential.

Assessment of Daily Dosage on FDA Drug Labels and DILI Severity

Numerous compound- and/or patient-specific risk factors can contribute to the susceptibility to DILI. IDILI has been shown to be dependent on both daily dose and extent of hepatic metabolism of a drug.[7,19,21,22] Lammert and coworkers[7,19] have attributed hepatic adverse events to compounds with significant hepatic metabolism and daily dose ≥ 50mg. Formation of reactive metabolites, high covalent body burden,[23,24] mitochondrial dysfunction (resulting in the depletion of cellular energy supply and the generation of damaging reactive oxygen species), cell damage from oxidative stress (caused by reactive oxygen or reactive nitrogen species), and local inflammatory effects.[25] All of these mechanisms are often interconnected and under various circumstances have been associated with the formation of chemically reactive metabolites.

We also previously evaluated the relationship between daily dosages ≥ 50mg against the already assessed FDA hepatic liability categories and DILI severity assessment.[8] Our analysis[15] concurs with the association of drugs being given at dosages ≥ 50mg/day having more adverse hepatic events. We have further evaluated this observation by examining the FDA hepatic liability distribution and DILI severity assessment. Drugs with a daily dose ≥ 50mg had a much higher frequency of toxicity as evidenced by the higher percentages in the "Warning and

Precautions", "Boxed Warning" and "Withdrawn" label sections (Figure 2A). For the DILI assessment in Figure 2B we also observe a higher frequency in DILI severity for compounds that are dosed at ≥ 50mg/day.

Although, there is strong evidence that dosages ≥ 50mg/day are associated with increased risk for hepatotoxicity, many drugs are safe at such dosages. For instance, the 50mg/day dosage cut off would predict that 44% of "No mention" and/or "No DILI" drugs (See Figure 2) exhibit "Not Safe" potential in terms of hepatotoxicity, supporting that daily dosage alone is not a reliable means of guiding the drug development process, regulatory application, and clinical practice.

Why should dose and lipophilicity be of predictive value? Likely this is because of the need for the liver to be exposed to a threshold level of the parent drug and/ or reactive metabolite. Lipophilic drugs are cleared by the liver and generally require biotransformation to be eliminated. As noted by the authors, a significant relationship was observed between the extent of hepatic

Why should dose and lipophilicity be of predictive value? Likely this is because of the need for the liver to be exposed to a threshold level of the parent drug and/ or reactive metabolite. Lipophilic drugs are cleared by the liver and generally require biotransformation to be eliminated. As noted by the authors, a significant relationship was observed between the extent of hepatic

BDDCS Classification Prior to Dosing in Humans

Although the finding of Uetrecht shows that idiosyncratic drug reactions were rare among individuals given drug doses <10mg/day and more likely among individuals given drug doses ≥ 1000mg/day,[26] the dose relationships can only be determined for a new molecular entity after the drug has been administered to human subjects/patients. In contrast, BDDCS class can be predicted prior to ever dosing the compounds to animals and humans as we have proposed

previously.[27] Hosey and Benet[28] showed that based on *in vitro* permeability measurements, the positive predictive value (PPV) for prediction of extensive metabolism were all 90% or greater. And recently Dave and Morris[29] showed that they were able to correctly predict highly soluble vs. poorly soluble drugs using measured solubility parameters with greater than 85% probability.

Drug Metabolism and Propagation of Drug Hypersensitivity Reactions

Drug metabolism also plays an important role in the initiation and propagation of drug hypersensitivity through the generation of neoantigens that are recognized by the cellular and humoral immune systems.[2] Although the majority of drug biotransformations occur in the liver, there is overwhelming evidence to suggest that localized drug metabolism by immune cells is critical for organ-specific reactions such as cutaneous adverse drug reactions.[2,30,31] These reactions are usually rare and are not typically present in animal species, but they can be serious and even fatal in humans[32,33] and may lead to the withdrawal of otherwise effective therapeutic agents. At present, during preclinical drug evaluation there are no widely accepted methods for the identification of drugs that may cause hypersensitivity or idiosyncratic reactions. It has been demonstrated that HLA-B*15:02 is not only a genetic marker but also a key determinant in the pathophysiology of carbamazepine related Stevens-Johnson Syndrome and Toxic Epidermal Necrolysis (SJS/TEN). We have previously assessed the use of the BDDCS to distinguish antiepileptic drugs (AEDs) associated with and without cutaneous adverse events by examining *in vitro* the binding relationship of AEDs to HLA-B*15:02 via surface plasmon resonance and clinical patient data from extensive reviews of medical records. We also evaluated the lack of benefit from a Hong Kong population policy on the effects of screening for HLA-B*15:02 and previous incorrect structure-activity hypotheses. Our analysis concludes that BDDCS Class 2

AEDs are more prone to cause adverse cutaneous reactions than certain BDDCS Class 1 AEDs and that BDDCS Class 3 drugs have the lowest levels of cutaneous adverse reactions. We proposed that BDDCS Class 3 AEDs should be preferentially used for patients with Asian backgrounds (i.e., Han Chinese, Thai and Malaysian populations) if possible and in patients predisposed to skin rashes.[14]

Alfirevic and Pirmohamed[34] and Urban *et al.*[35] have summarized the current state of pharmacogenomics and suggested that although certain HLA and other differences are related to a higher susceptibility of DILI from a number of agents, the actual number of drugs identified as having these genetic risks is still quite small, and the accuracy of most polymorphisms is limited. Although significant advances in our hepatotoxicity knowledge base have been made by the DILI Network and others,[36] when it comes to identifying the specific components of DILI risk, it appears to be much more complex than just being a matter of daily dose or drug disposition.[21]

Statistical Tests of DILI Predictivity

Most frequently statistical measures of the performance of binary classification tests are reported as Sensitivity and Specificity. We believe that DILI predictivity deserves a more stipulated analysis. In bringing a new molecular entity (NME) to market, predictive statistical measures for DILI should address a) how good is the metric in correctly identifying compounds that should be avoided in terms of any further experimental work early in drug development? And b) what are the chances that an NME showing few preclinical predictors of DILI, will in fact elicit DILI after the company has spent considerable time and money in bringing the drug to market? We believe that the “a” question is best evaluated by the positive predicted value (PPV),

the percentage of positive predictions of DILI that actually cause DILI (i.e., True Positive divided by the sum of True Positive and False Positive). If the PPV value for a DILI metric is very high, companies will correctly suspend development of DILI causing compounds very early in the process, and there is only a small chance that a compound not causing DILI will be discarded. We prefer this PPV measure to Specificity (i.e., True Negative divided by the sum of True Negative and False Positive) since we believe PPV more correctly addresses the early issue of stopping development, where Specificity relates to final outcome.

The "b" question is best evaluated by the false negative rate (FNR), the percentage of NMEs that a sponsor takes through extensive and expensive human studies and the regulatory approval process, where DILI manifests after the drug goes on the market. A good DILI predictive metric should show a low FNR (calculated as False Negative divided by False Negative + True Positive). Note the FNR percentage is just 100 minus the Sensitivity percentage. We also list the accuracy (ACC) of the various methods in our tabular comparisons (i.e., Total Positive + Total Negative divided by number of compounds tested). A high PPV metric saves money early in the process by not pursuing compounds highly suspected to cause DILI, but a low FNR metric is probably more important, potentially avoiding very large sunk costs and human suffering after considerable work.

Comparison of *In Vitro* Mechanism Based Toxicity Endpoints

Although, a number of compound–specific liability factors have been linked with DILI susceptibility, it is difficult to understand which risk factors are more important in patient-specific responses and/or environmental stimuli. One approach followed by many research

groups is to assess and reduce some of the more common, drug-specific factors in a set of targeted *in vitro* assays. The most common mechanisms covered in *in vitro* high throughput screening assays include reactive metabolite formation and covalent binding,[37,38] inhibition of drug transporters involved in hepatobiliary elimination of bile acids and other metabolic endogenous products (BSEP, MRPs),[5,39] mitochondrial toxicity[40] and different cellular toxicity assays covering the formation of drug-metabolites.[6,41–44] Various approaches are used in the pharmaceutical industry for hazard identification and risk assessment of reactive metabolites and more integrated strategies that include measures of the initial mechanism of toxicity have been highlighted in our analysis.

We have previously performed a comparison of the different predictive metrics in the various assays measuring key mechanisms of toxicity endpoints associated with DILI from the Schadt *et al.* data set.[43] The toxicity endpoints were monitored in a panel consisting of assays assessing the generation of reactive metabolites tested via GSH adduct formation, P450 3A4 time-dependent inhibition (TDI), BSEP inhibition, mitochondrial toxicity and cytotoxicity. In the Schadt *et al.* data set of 120 marketed compounds, 14 compounds had not been BDDCS classified. Our analysis is depicted in Table 1. (Supplemental Table S1 includes Sensitivity and Specificity statistical analyses.) The assays that performed the best were GSH adduct formation and BSEP inhibition. We noted that although the positive predictive value (PPV) for these measurements were somewhat better than for BDDCS Class 2 classification, the false negative rate (FNR) for these measures was much greater than BDDCS Class 2, so that in terms of accuracy (ACC), the GSH and BSEP assays were no better than just avoiding BDDCS Class 2 drugs. When GSH and BSEP assays were combined with BDDCS Class 2, again higher PPV values are obtained, but because FNR also increased, ACC is not better than just avoiding

BDDCS Class 2 drugs. A slightly higher ACC is obtained when all of the mechanisms of toxicity endpoints are confirmed, due to the low FNR. However, having a PPV of only 65.1% does not give much confidence. We have observed better predictability values for the correlation with MRP3 and MRP4 inhibition, with MRP3 or MRP4 inhibitors giving the best predictability.[16] The basolateral efflux transporters, MRP3 and MPR4, play a minor role in bile acid efflux under normal conditions, but they are up-regulated under cholestatic conditions to compensate for impaired biliary excretion. Thus, impaired function of MRP3 and MRP4 by drugs, or genetic polymorphisms resulting in reduced-function variants may result in accumulation of toxic bile acids in hepatocytes.[45] But further studies are needed in order to investigate whether this could be a biomarker of value for DILI screening.

Although there may be some general trends between simple physical parameters, it is unlikely that such considerations could accurately predict risk. This problem could potentially be alleviated by the new *in vitro* approaches in physiological test systems with model hepatotoxins and utilization of state of the art instrumentation currently being evaluated encompassing chemical and biological factors associated with hepatotoxicity earlier in drug development, as well as the commercial emergence of long-enduring primary hepatocyte-based cultures.[15,46,47]

Assessment of BDDCS Classification on BSEP Inhibition and DILI Risk

The accumulation of bile acids within hepatocytes is thought to be a primary mechanism for the development of DILI, although as we show with the Schadt *et al.* data, this is not confirmed. Inhibition of the bile salt export pump (BSEP) by a drug has been implicated as a risk factor for a drug's potential to cause DILI. However, few reports indicate that drug-induced

BSEP dysfunction actually leads to hepatotoxicity, and the relationship between drug-induced BSEP dysfunction and liver injury risk is yet to be determined. Recently, the International Transporter Consortium has highlighted BSEP as one of the emerging transporters that need to be considered when evaluating drug safety. However, the practical utility of this approach still needs to be further evaluated. We analyzed further data encompassing the relationship between a compound's ability to inhibit BSEP function *in vitro* and cause liver injury in humans *in vivo* using a compilation of published DILI datasets that have screened for BSEP inhibitors, other hepatic transporters and other mechanism based toxicity endpoints such as the mitochondrial toxicity assay.[6,43,48,49] We evaluated the information provided by using BDDCS in order to understand the inhibition propensity of BSEP. Our results demonstrate that there is little support for *in vitro* BSEP inhibition being usefully DILI predictive. Rather we show that the most potent BSEP inhibitors are BDDCS Class 2 drugs, which we have demonstrated previously is the BDDCS class most likely to be DILI related.[16]

When BSEP inhibition data by Pedersen *et al.*[50] were correlated with the Chen DILI Assessment,[51] we observed no discernible pattern[16] (See Figure 3A). For the BDDCS classification, we observe that the great majority of strong BSEP inhibitors are BDDCS Class 2 drugs, with concomitant decreases in the percentages of BDDCS class 1 and 3 drugs as BSEP inhibition increases, as depicted by Figure 3B. These results along with our analyses of other datasets,[16] show that *in vitro* measures of BSEP inhibition by small molecules is not a strong susceptibility factor. It is unclear as to what extent BSEP inhibition is functionally significant *in vivo*. Furthermore, our analyses demonstrate that the great majority of compounds that have been associated with DILI and show strong BSEP inhibition potential *in vitro* are also BDDCS Class 2 drugs. Because we are able to make similar predictions based on BDDCS determinant

characteristics, this leads us to discount the predictive ability of *in vitro* BSEP inhibition for DILI. We have previously observed that as hepatic warning severity increases, the proportion of BDDCS Class 2 drugs increases and the proportions of both BDDCS Class 1 and 3 drugs decrease.[15] We conclude that previous analyses predicting that BSEP inhibition leads to DILI may have been confounded by the observation that most BSEP inhibitors are BDDCS Class 2 drugs, which show a high prevalence for DILI.[16]

Aleo *et al.*[6] suggest that mitochondrial toxicity together with BSEP inhibition may provide improved DILI predictability. When we analyzed the predictability of BSEP inhibition together with mitochondrial toxicity, we observe that BDDCS class 2 characterization shows comparable results.[16] This is further confirmed in the Schadt *et al.* data set where we show in Table 1 that combining BSEP inhibition and mitotoxicity yields a very high FNR and no improvement in ACC. Thus, we believe that neither BSEP inhibition nor mitochondrial toxicity are useful independent or combined predictors of DILI.

Studies with Long-Enduring Metabolically Competent Hepatocyte Co-Cultures

Many of the deficiencies of the high-throughput DILI screens employed by the pharmaceutical industry based on *in vitro* measures of hepatocyte toxicity theoretically can be overcome with long-enduring metabolically competent hepatocyte co-cultures. The long-enduring status has the potential to allow formation of active metabolites and/or accumulation of intracellular toxic substances in hepatocytes due to inhibition of relevant efflux transporters, as well as giving sufficient time for the hepatotoxicity to manifest. A December 2017 study[47] describes results for 19 identified drugs (12 known to exhibit clinical DILI and 7 that were negative), where measures of TC_{50}/C_{max} were obtained in human hepatocyte mono-cultures after

1 day and co-cultures after 1, 7 and 14 days. The results compared to BDDCS classification are summarized in Table 2.

None of the seven DILI negative drugs (propranolol, rosiglitazone, diphenhydramine, isoproterenol, kanamycin, macitentan and primidone) gave a positive reading; therefore, all PPV values were 100%. The two clinical DILI drugs (bicalutamide and tacrine) that gave no positive readings under any conditions exhibited decreased TC_{50}/C_{max} values with time but did not reach toxic readings. It is interesting to note that BDDCS Class 2 categorization approached the day 1 accuracy values, as seen in the comparisons in Table 2 and our more extensive analyses.[15,16] As we have noted previously, [14–16] publication of toxicity predictors concentrate in achieving high PPV rates. In contrast, in drug development what is critical to the industry in terms of cost and effort is a predictive metric with a low FNR. In Table 2 we also list the results for our analysis using the criteria of clinical dose ≤ 50 mg as proposed by Lammert *et al.*,[7] which yielded the lowest FNR of all the data sets we analyzed,[15,16] for the 19 drugs investigated by Novik *et al.*[47] and the Chen *et al.*[8] DILI data set for BDDCS classifiable drugs. However, it must be remembered that BDDCS classification and the co-culture analysis reported by Novik *et al.*[47] can occur very early in drug development, certainly before knowledge of the clinical dose and C_{max} is obtained. The downside to the Novik et al.[47] analysis, based on the ratio of TC_{50} to C_{max}, and the criteria of clinical dose ≤ 50 mg is that these values will not be known early in drug development.

Yet, even with this limitation, we are encouraged that an additional predictive measure of DILI (in terms of time of incubation) is available. We suspect that if the highly positive results reported by Novik *et al.*[47] for 19 drugs can be confirmed in further studies of many more therapeutic agents, this information will be of great value to drug sponsors and regulators in

evaluating Phase 3 results and NDA applications. Furthermore, we expect that positive results from long-enduring hepatocytes in predicting DILI may be combined with advances utilizing *in silico* predictors of pharmacokinetics and translation of *in vitro* measures of effective concentrations employing novel biomarker systems facilitating future DILI hypothesis measures that can be shown to be superior to the BDDCS classification as demonstrated in Table 2.

Studies with 3D Spherical Human Liver Microtissues (hLiMT) and 2D Primary Human Hepatocytes (PHH)

3D Spherical Human Liver Microtissues (hLiMT) have emerged as promising tools to assess mechanisms of hepatotoxicity, as they demonstrate enhanced liver phenotype, metabolic activity, and stability in culture not attainable with conventional 2D hepatic models. hLiMT are made up of primary hepatocytes and non-parenchymal cells that can be used for repeat-dose long term treatment. Proctor and co-workers investigated the cytotoxicity of 110 marketed drugs in 3D Spherical Human Liver Microtissues (hLiMT) and in 2D primary hepatocytes (PHH) from the same human donor.[52] To complement our analysis, we performed a BDDCS classification and evaluated the prediction of these two metabolic-competent hepatocyte cultures on the 100 drugs (of the 110) that had been BDDCS classified. Proctor *et al.*[52] evaluated various thresholds of IC_{50} among 10, 25, 50 and 100μM in terms of DILI predictivity as well as margin of safety measures (IC_{50}/C_{max}) at the same thresholds for the 3D hLiMT and 2D PHH analyses. In Table 3 we reproduce here the results for the 100 μM threshold for the 100 BDDCS classified compounds using drug-induced cytotoxicity as an endpoint since this threshold exhibited the highest accuracy (the results for all four thresholds together with Sensitivity and Specificity values are given Supplementary Table S2 for all 110 compounds and in Table S3 for the 100 BDDCS

classified compounds). The 3D hLiMT 14 day measures yield lower FNR percentages and higher ACC than the 2D PHH 2 day measures and BDDCS class 2 assignment. However, as seen in Table 3, the best differentiating measurement to predict DILI toxicity is seen for C_{max} > 1.3μM, a parameter that will be influenced by dose size. Although, Proctor and co-workers conclude that the hLiMT demonstrate sufficient capability to warrant exploratory liver injury biomarker investigation, we remain skeptical about the usage of the hLiMT to screen for DILI given that we did not see that any of the proposed criteria had a better prediction than C_{max} > 1.3μM in this large, 100 drug dataset.

Why are BDDCS Class 2 Drugs More Toxic than BDDCS Class 1 Drugs?

Several studies have also shown a correlation of total administered dose alone or in combination with drug lipophilicity with higher risk of DILI. However, as we show in Fig. 2 and in our discussion above, dose alone is not able to accurately discriminate all drugs causing DILI. Chen *et al.*[8] proposed a Rule of 2 where PPV for DILI was very high when considering drugs with cLogP ≥3.0 (calculated lipid water partition coefficient) and dose > 100mg. However, we have shown that the Rule of 2 was slight less accurate than just BDDCS Class 2 assignment.[15] But, using the Lammert *et al*.[7] cutoff of dose > 50 mg being more likely to cause DILI, we note that from our compilation of the highest approved dose strength[53], 54% of Class 2 drugs exceed 50 mg, while only 33% of Class 1 drugs do so.

Highly lipophilic drugs are cleared by the liver and generally require biotransformation to be eliminated. The parameter cLogP may simply be a surrogate for extensive biotransformation and hepatic exposure to a reactive metabolite. If cLogP could differentiate DILI potential, we would see equal chances of BDDCS Class 1 and 2 drugs leading to DILI toxicity. As seen in

Figs. 1, BDDCS Class 2 compounds predominate among the most severe hepatic toxicities. We also observe a clear differentiation between BDDCS Class 2 and 1 in terms of DILI predictability in Table 1. Furthermore, in our previous analysis,[15] we have observed that PPV for cLogP ≥ 3 alone is fairly low (76.1% and ACC is 52.3%). Therefore, we do not believe extensive metabolism is an adequate DILI predictor.

A major finding in the development of the BDDCS was the recognition that BDDCS Class 1 drugs, i.e. extensively metabolized, highly permeable, highly soluble, may be shown *in vitro* to be substrates of both uptake and efflux transporters, but that effects of transporters on BDDCS Class 1 drugs are essentially clinically insignificant in the liver and intestine, as well as the brain. Thus, the unbound concentrations of BDDCS Class 1 drugs in the systemic circulation will reflect unbound concentrations in the liver as well as in the rest of the body, since it is transporters that lead to differences in unbound concentrations in different organs. According to BDDCS[53,54] approximately 40% of marketed drugs (i.e., those that are Class 1) will still follow the equivalent free drug concentration hypothesis. However, this will not be true for BDDCS Classes 2, 3 and 4 drugs where transporter effects may lead to different unbound concentrations in the liver and throughout the body. That is, Class 1 drugs will follow the long held assumption in deriving pharmacologic/toxicologic relationships that free drug concentrations are the same throughout the body. But this assumption in pharmacology was made prior to any recognition of the importance of drug transporters in controlling permeability.

It is important to recognize that the compounds evaluated here are drugs that reach the market where sponsors were able to convince the regulatory agencies based on *in vitro* and preclinical animal studies that toxicity potential, particularly DILI, would be manageable or at least acceptable when the drugs reached the market and were taken by large patient populations

as compared to those limited number of patients studied during drug development. Thus, according to our hypothesis, drug company sponsors in their preclinical and clinical studies of Class 1 drugs would be able to reasonably predict drug concentrations in the liver and throughout the body. In contrast, for BDDCS Class 2 drugs, where metabolism is the significant process of elimination, drug concentration measurements in the systemic circulation for these compounds both in the preclinical and clinical studies may poorly predict what concentrations are present in the liver and in other organs of the body. And since it is obvious that DILI occurs more frequently with metabolized drugs, studies in drug development with Class 2 drugs would be poorer predictors of toxicity potential due to the challenges to estimate intracellular concentrations and metabolic processes. Thus, the prevalence of DILI with BDDCS Class 2 drugs could just be circumstantial in that sponsors would be unable to properly evaluate hepatic toxicity for these compounds in designing their clinical studies. This problem could potentially be alleviated by new *in vitro* approaches and utilization of state of the art instrumentation currently being evaluated.

Conclusions

The application of the BDDCS methodology can help evaluate the potential validity of risk assessment hypotheses. The BDDCS Class 2 susceptibility factor yields similar and in a number of cases better accuracy than the DILI predictive potential biomarkers of other methodologies. Since there is no mechanistic basis for BDDCS Class 2 drugs being most DILI related, if an alternate hypothesis is no more predictive than BDDCS Class assignment, we maintain that the alternate hypothesis is not sufficiently predictive, nor a mechanistic valid hypothesis. As seen in Fig. 1, the BDDCS Class 2 versus Class 1 differentiation only becomes

evident with the most severe hepatic toxicities, and then only a 2:1 differentiation between BDDCS Class 2 versus Class 1 is found. Lammert *et al.*'s[19] assertion that extensive metabolized compounds are at higher risk of developing DILI can be much improved by differentiating BDDCS Class 2 from BDDCS Class 1 drugs. Daily dosage ≥ 50mg alone can only depict a clear relationship with dose with compounds that have been previously associated with DILI, but very limited predictability in differentiating compounds with "No DILI" assignment. There is a general acceptance that BSEP inhibition is a source of toxicity. However, according to our analysis of DILI this is not true for *in vitro* measures of BSEP inhibition. What we find is that most DILI occurs with BDDCS Class 2 compounds and almost all BSEP substrates and inhibitors are Class 2 compounds, but we observe no relationship between the strength of *in vitro* BSEP inhibition and toxicity, which makes us believe that the generally held hypothesis is incorrect.

Our review of the BDDCS analysis alongside other DILI toxicity potential biomarkers show that none of the current *in vitro* methodologies are sufficiently accurate and effective in allowing early identification of new molecular entities that will be DILI free. The comparison of proposed DILI predictive methodology with BDDCS assignment offers a useful tool by which new DILI predictive hypotheses can be evaluated. Thus, we are encouraged by the apparent increased sensitivity, as compared to BDDCS Class 2, for long-enduring metabolically competent hepatocyte co-cultures. Furthermore, using BDDCS classification and finding that a compound is Class 2, one would recognize that priority should be given to more aggressively investigate its DILI potential in mechanistic DILI assays. Some DILI risk factors can be mitigated during the drug design/development process to identify drugs with better chemical attributes with reduced potential to cause human DILI. Hopefully, development of mechanism

based toxicity endpoints, such as those previously proposed by Chen *et al*.[51] and Schadt *et al.*,[43] as discussed above, will improve future predictability.

Our review of this work has clearly pointed out that many of the published "predictive DILI" hypotheses do no better than just avoiding BDDCS Class 2 drugs. We propose that comparison of predictive DILI hypotheses with BDDCS class assignment is a useful exercise in determining the relevance of predictive metrics. The results presented herein illustrate how BDDCS can be applied to better understand clinically observed hepatotoxicity and aid in the DILI risk assessment of new molecular entities.

Why should dose and lipophilicity be of predictive value? Likely this is because of the need for the liver to be exposed to a threshold level of the parent drug and/or reactive metabolite. Lipophilic drugs are cleared by the liver and generally require biotransformation to be eliminated. As noted by the authors, a significant relationship was observed between the extent of hepatic

Funding

Rosa Chan was supported by the American Association of Pharmaceutical Scientists Graduate Fellowship, American Foundation for Pharmaceutical Education Pre-Doctoral Fellowship, North American Graduate Fellowship from the American College of Toxicology, Supplemental Training for Education Program from the Society of Toxicology, and NIGMS grant R25 GM56847. Dr. Benet is a member of the UCSF Liver Center supported by NIH Grant P30 DK026743.

Conflicts

Dr. Benet is a founder, holds stock and is chairman of the scientific advisory board of Hμrel Corporation. Dr. Chan declares no conflicts or competing financial interest.

References

1 F. Shah, L. Leung, H. A. Barton, Y. Will, A. D. Rodrigues, N. Greene and M. D. Aleo, Setting clinical exposure levels of concern for drug-induced liver injury (DILI) using mechanistic in vitro assays, *Toxicol. Sci.*, 2015, **147**, 500–514.

2 R. A. Thompson, E. M. Isin, M. O. Ogese, J. T. Mettetal and D. P. Williams, Reactive metabolites: current and emerging risk and hazard assessments, *Chem. Res. Toxicol.*, 2016, **29**, 505–533.

3 N. Kaplowitz, Avoiding idiosyncratic DILI: two is better than one, *Hepatology*, 2013, **58**, 15–17.

4 R. E. Morgan, M. Trauner, C. J. van Staden, P. H. Lee, B. Ramachandran, M. Eschenberg, C. a. Afshari, C. W. Qualls, R. Lightfoot-Dunn and H. K. Hamadeh, Interference with bile salt export pump function is a susceptibility factor for human liver injury in drug development, *Toxicol. Sci.*, 2010, **118**, 485–500.

5 S. Dawson, S. Stahl, N. Paul, J. Barber and J. G. Kenna, In vitro inhibition of the bile salt export pump correlates with risk of cholestatic drug-induced liver injury in humans, *Drug Metab. Dispos.*, 2012, **40**, 130–138.

6 M. D. Aleo, Y. Luo, R. Swiss, P. D. Bonin, D. M. Potter and Y. Will, Human drug-induced liver injury severity is highly associated with dual inhibition of liver mitochondrial function and bile salt export pump, *Hepatology*, 2014, **60**, 1015–1022.

7 C. Lammert, S. Einarsson, C. Saha, A. Niklasson, E. Bjornsson and N. Chalasani, Relationship between daily dose of oral medications and idiosyncratic drug-induced liver injury: search for signals, *Hepatology*, 2008, **47**, 2003–2009.

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

9 J. H. Hoofnagle, J. Serrano, J. E. Knoben and V. J. Navarro, LiverTox: a website on drug induced liver injury, *Hepatology*, 2013, **57**, 873–874.

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

11 M. Chen, H. Bisgin, L. Tong, H. Hong, H. Fang, J. Borlak and W. Tong, Toward predictive models for drug-induced liver injury in humans : are we there yet ?, *Biomark. Med.*, 2014, **8**, 201–213.

12 M. P. Gleeson, Generation of a set of simple, interpretable ADMET rules of thumb, *J. Med. Chem.*, 2008, **51**, 817–834.

13 L. Z. Benet, Predicting drug disposition via application of a biopharmaceutics drug disposition classification system, *Basic Clin. Pharmacol. Toxicol.*, 2010, **106**, 162–167.

14 R. Chan, C. Wei, Y. Chen and L. Z. Benet, Use of the biopharmaceutics drug disposition classification system (BDDCS) to help predict the occurrence of idiosyncratic cutaneous adverse drug reactions associated with antiepileptic drug usage, *AAPS J.*, 2016, **18**, 757–766.

15 R. Chan and L. Z. Benet, Evaluation of DILI predictive hypotheses in early drug development, *Chem. Res. Toxicol.*, 2017, **30**, 1017–1029.

16 R. Chan and L. Z. Benet, Measures of BSEP inhibition in vitro are not useful predictors of DILI, *Toxicol. Sci.*, 2018, **162**, 499-508.

17 C.-Y. Wu and L. Z. Benet, Predicting drug disposition via application of BCS: transport/absorption/elimination interplay and development of a Biopharmaceutics Drug Disposition Classification System, *Pharm. Res.*, 2005, **22**, 11–23.

18 A. E. Trontell, How the US food and drug administration defines and detects adverse drug events, *Curr. Ther. Res.*, 2001, **62**, 641–649.

19 C. Lammert, E. Bjornsson, A. Niklasson and N. Chalasani, Oral medications with significant hepatic metabolism at higher risk for hepatic adverse events, *Hepatology*, 2010, **51**, 615–620.

20 P. Barton and R. J. Riley, A new paradigm for navigating compound property related drug attrition, *Drug Discov. Today*, 2016, **21**, 72–81.

21 J. H. Lewis, Drug-induced liver injury, dosage, and drug disposition: is idiosyncrasy really unpredictable?, *Clin. Gastroenterol. Hepatol.*, 2014, **12**, 1556–1561.

22 M. Chen, A. Suzuki, J. Borlak, R. J. Andrade and M. Isabel Lucena, Drug-induced liver injury: interactions between drug properties and host factors, *J. Hepatol.*, 2015, **63**, 503–514.

23 A. Srivastava, J. L. Maggs, D. J. Antoine, D. P. Williams, D. A. Smith and B. K. Park, Role of reactive metabolites in drug-induced hepatotoxicity, *Handb. Exp. Pharmacol.*, 2010, **196**, 165–194.

24 S. R. Knowles, J. Uetrecht and N. H. Shear, Idiosyncratic drug reactions: the reactive metabolite syndromes., *Lancet*, 2000, **356**, 1587–1591.

25 H. Jaeschke, G. J. Gores, A. I. Cederbaum, J. A. Hinson, D. Pessayre and J. J. Lemasters, Mechanisms of hepatotoxicity, *Toxicol. Sci.*, 2002, **65**, 166–176.

26 J. Uetrecht and L. Dan, Idiosyncratic drug reactions: current understanding, *Annu. Rev. Pharmacol. Toxicol*, 2007, **47**, 513–539.

27 F. Broccatelli, G. Cruciani, L. Z. Benet and T. I. Oprea, BDDCS class prediction for new molecular entities, *Mol. Pharm.*, 2012, **9**, 570–580.

28 C. M. Hosey and L. Z. Benet, Predicting the extent of metabolism using in vitro permeability rate measurements and in silico permeability rate predictions, *Mol. Pharm.*, 2015, **12**, 1456–1466.

29 R. A. Dave and M. E. Morris, Novel high/low solubility classification methods for new molecular entities, *Int. J. Pharm.*, 2016, **511**, 111–126.

30 C. Ju and J. P. Uetrecht, Detection of 2-hydroxyiminostilbene in the urine of patients taking carbamazepine and its oxidation to a reactive iminoquinone intermediate., *J. Pharmacol. Exp. Ther.*, 1999, **288**, 51–56.

31 J. M. Baron, D. Höller, R. Schiffer, S. Frankenberg, M. Neis, H. F. Merk and F. K. Jugert, Expression of multiple cytochrome p450 enzymes and multidrug resistance-associated transport proteins in human skin keratinocytes, *J. Invest. Dermatol.*, 2001, **116**, 541–8.

32 O. Pelkonen and H. Raunio, Metabolic activation of toxins: tissue-specific expression and metabolism in target organs, *Environ. Health Perspect.*, 1997, **105**, 767–774.

33 J. P. Uetrecht, The role of leukocyte-generated reactive metabolites in the pathogenesis of idiosyncratic drug reactions, *Drug Metab. Rev.*, 1992, **24**, 299–366.

34 A. Alfirevic and M. Pirmohamed, Predictive genetic testing for drug-induced liver injury: considerations of clinical utility, *Clin. Pharmacol. Ther.*, 2012, **92**, 376–380.

35 T. J. Urban, Y. Shen, A. Stolz, N. Chalasani, R. J. Fontana, J. Rochon, D. Ge, K. V. Shianna, A. K. Daly, M. I. Lucena, M. R. Nelson, M. Molokhia, G. P. Aithal, A. Floratos, I. Pe'er, J. Serrano, H. Bonkovsky, T. J. Davern, W. M. Lee, V. J. Navarro, J. A. Talwalkar, D. B. Goldstein and P. B. Watkin, Limited contribution of common genetic variants to risk for liver injury due to a variety of drugs, *Pharmacogenet. Genomics*, 2013, **22**, 784–795.

36 N. Chalasani, R. J. Fontana, H. L. Bonkovsky, P. B. Watkins, T. Davern, J. Serrano, H. Yang and J. Rochon, Causes, clinical features, and outcomes from a prospective study of drug-induced liver injury in the United States, *Gastroenterology*, 2008, **135**, 1924–1934.

37 S. Nakayama, R. Atsumi, H. Takakusa, Y. Kobayashi, A. Kurihara, Y. Nagai, D. Nakai and O. Okazaki, A zone classification system for risk assessment of idiosyncratic drug toxicity using daily dose and covalent binding, *Drug Metab. Dispos.*, 2009, **37**, 1970–1977.

38 T. Usui, M. Mise, T. Hashizume, M. Yabuki and S. Komuro, Evaluation of the potential for drug-induced liver injury based on in vitro covalent binding to human liver proteins, *Pharmacology*, 2009, **37**, 2383–2392.

39 R. E. Morgan, C. J. van Staden, Y. Chen, N. Kalyanaraman, J. Kalanzi, R. T. Dunn, C. A. Afshari and H. K. Hamadeh, A multifactorial approach to hepatobiliary transporter assessment enables improved therapeutic compound development, *Toxicol. Sci.*, 2013, **136**, 216–241.

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

41 Y.-M. Wang, S. C. Chai, C. T. Brewer and T. Chen, Pregnane X receptor and drug-induced liver injury, *Expert Opin. Drug Metab. Toxicol.*, 2014, **10**, 1521–1532.

42 J. J. Xu, P. V. Henstock, M. C. Dunn, A. R. Smith, J. R. Chabot and D. de Graaf, Cellular imaging predictions of clinical drug-induced liver injury, *Toxicol. Sci.*, 2008, **105**, 97–105.

43 S. Schadt, S. Simon, S. Kustermann, F. Boess, C. McGinnis, A. Brink, R. Lieven, S. Fowler, K. Youdim, M. Ullah, M. Marschmann, C. Zihlmann, Y. M. Siegrist, A. C. Cascais, E. Di Lenarda, E. Durr, N. Schaub, X. Ang, V. Starke, T. Singer, R. Alvarez-Sanchez, A. B. Roth, F. Schuler and C. Funk, Minimizing DILI risk in drug discovery - a screening tool for drug candidates, *Toxicol. Vitr.*, 2015, **30**, 429–437.

44 S. Russmann, G. A. Kullak-Ublick and I. Grattagliano, Current concepts of mechanisms in drug-induced hepatotoxicity, *Curr. Med. Chem.*, 2009, **16**, 3041–3053.

45 K. Yang, K. Köck, A. Sedykh, A. Tropsha and K. L. R. Brouwer, An updated review on drug-induced cholestasis: mechanisms and investigation of physicochemical properties and pharmacokinetic parameters, *J. Pharm. Sci.*, 2013, **102**, 3037–3057.

46 E. L. LeCluyse, R. P. Witek, M. E. Andersen and M. J. Powers, Organotypic liver culture models: meeting current challenges in toxicity testing, *Crit. Rev. Toxicol.*, 2012, **42**, 501–48.

47 E. I. Novik, J. Dwyer, J. K. Morelli, A. Parekh, C. Cho, E. Pludwinski, A. Shrirao, R. M. Freedman, J. S. MacDonald and Z. Jayyosi, Long-enduring primary hepatocyte-based co-cultures improve prediction of hepatotoxicity, *Toxicol. Appl. Pharmacol.*, 2017, **336**, 20–30.

48 M. D. Aleo, F. Shah, K. He, P. D. Bonin and A. D. Rodrigues, Evaluating the role of multidrug resistance protein 3 (MDR3) inhibition in predicting drug induced liver injury using 125 pharmaceuticals, *Chem. Res. Toxicol.*, 2017, **30**, 1219–1229.

49 K. Köck, B. C. Ferslew, I. Netterberg, K. Yang, T. J. Urban, P. W. Swaan, P. W. Stewart and K. L. Brouwer, Risk factors for development of cholestatic drug-induced liver injury: inhibition of hepatic basolateral bile acid transporters MRP3 and MRP4, *Drug Metab. Dispos.*, 2014, **42**, 665–674.

50 J. M. Pedersen, P. Matsson, C. A. S. Bergström, J. Hoogstraate, A. Norén, E. L. LeCluyse and P. Artursson, Early identification of clinically relevant drug interactions with the human bile salt export pump (BSEP/ABCB11), *Toxicol. Sci.*, 2013, **136**, 328–343.

51 M. Chen, J. Borlak and W. Tong, A Model to predict severity of drug-induced liver injury in humans, *Hepatology*, 2016, **64**, 931–940.

52 W. R. Proctor, A. J. Foster, J. Vogt, C. Summers, B. Middleton, M. A. Pilling, D. Shienson, M. Kijanska, S. Ströbel, J. M. Kelm, P. Morgan, S. Messner and D. Williams, Utility of spherical human liver microtissues for prediction of clinical drug-induced liver injury, *Arch. Toxicol.*, 2017, **91**, 2849–2863.

53 L. Z. Benet, F. Broccatelli and T. I. Oprea, BDDCS applied to over 900 drugs, *AAPS J.*, 2011, **13**, 519–547.

54 L. Z. Benet, The role of BCS (Biopharmaceutics Classification System) and BDDCS (Biopharmaceutics Drug Disposition Classification System) in drug development, *J. Pharm. Sci.*, 2013, **102**, 34–42.