ELSEVIER

The future of drug safety testing: expanding the view and narrowing the focus

James L. Stevens and Thomas K. Baker

Division of Toxicology, Lilly Research Laboratories, Lilly Corporate Center, Indianapolis, IN 46285, United States

Drug safety remains a high profile issue at a time when the cost and time required to develop a new drug are at an all time high. Balancing risk against the expected clinical benefit is the primary purpose of preclinical and clinical testing. We offer an expanded view on the application of predictive strategies and technologies to early safety decisions and suggestions to narrow the focus for improving preclinical safety testing to the problems that contribute most to adverse drug reactions.

Introduction

Pressure to improve drug safety is increasing. This occurs at a time when industry productivity is at an all time low while the cost and time required to develop a new drug are at an all time high [1]. Toxicity is not only a safety concern but also a major cause of compound attrition during drug development. Accurately predicting which molecules are safe and less likely to encounter development hurdles can both protect patients and improve productivity.

Minimizing risk to people is essential, but eliminating risk is neither possible nor practical. Balancing predicted risk against expected clinical benefit is the essence of preclinical and clinical testing. The current preclinical testing paradigm, established over 30 years ago, has improved drug safety markedly; evidence suggests that 70% of human toxicity seen during clinical trials is predicted by preclinical studies [2]. This may be an underestimate because many compounds are terminated for safety concerns before reaching the clinic [3]. Thus, targeting improvements to the existing paradigm in the near term may be more useful than a complete overhaul. Identifying the areas to target is the key question.

On an average, it takes at least ten years to deliver a new drug to patients [4]. Much of this time is spent testing the efficacy and safety of a single chemical structure. We offer a perspective on improving drug safety through better compound selection. The focus is on small molecules, but some principles may apply to bioproducts. Our intent is to expand the view on how early decisions impact later development and narrow the focus on solving problems that contribute most to adverse drug reactions. We address strategies to improve drug safety in human through earlier preclinical intervention. The primary focus will be preventable nonidiosyncratic adverse reactions, sometimes defined as dose-dependent or intrinsic toxicity (see [5,6] for additional references). We will also avoid the discussion of whether or not the so-called idiosyncratic toxicity is truly idiosyncratic or could have been predicted from preclinical data [7].

Narrowing the focus: What are the most important problems in the near term?

Available studies suggest that cardiac and hepatic toxicity contribute disproportionately to drug withdrawals. Drawing from reviews [8–10], and the CDER Reports to the Nation [11] we identified 47 drugs withdrawn from the market (Fig. 1). Of these, 15 were terminated for hepatoxocity and 21 were withdrawn for cardiac safety. Among those terminated for cardiac safety, 11 were for *torsades de pointe*. Others report similar conclusions, although the relative contribution of heart and liver is higher in our analysis, perhaps owing to the inclusion of more compounds [12] or exclusion of drugs withdrawn for reasons other than human toxicity [9].

Adverse drug reactions also occur during clinical testing, before approval. Published data are scarce, but a retrospective analysis determined the concordance between clinical and preclinical toxicity [2]. Of the clinical adverse events reported, 71% were contributed by hepatic (14%), cardiovascular (16%), neurological (22%) and gastrointestinal (19%) toxicity (see Table 3 in Olson

Corresponding author: Baker, Stevens, J.L. (stevens_james_l@lilly.com), T.K. (baker_thomas_k@lilly.com)

1359-6446/06/$ - see front matter doi:10.1016/j.drudis.2008.11.009

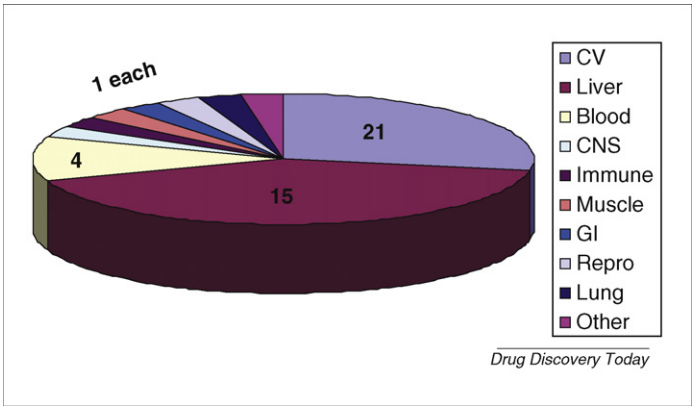

FIGURE 1

Summary of target organ contributions to drug withdrawals 1975–2007. Data were compiled from reviews [8–10] and CDER Reports to the Nation [11] as well as the CDER website. Target organs were identified on the basis of the reasons for withdrawal as noted in the references.

et al. [2]). Notably, nonrodent studies were more predictive than rodent studies. Hepatic, cardiac and neurological toxicity caused 66% of the terminations during clinical testing. When preclinical studies accurately predicted human toxicity, 94% were detected in studies 30 days or less, those studies that generally support Phase I trials.

This analysis leads to three conclusions. First, reducing three primary causes of toxicity (hepatic, cardiac and neurological) would have the highest impact on reducing risk during clinical trials. Second, a twofold reduction in drug withdrawals for cardiac or liver safety concerns would decrease withdrawals overall by as much as 40%. Third, efforts to improve preclinical testing should focus on studies of 30-day duration or less to have the highest impact.

Preclinical focus: advancing safer compounds into clinical testing

Recent reviews summarize the preclinical testing paradigm [3,13]. Preclinical testing is designed to determine; 1 – the dose limiting toxicity (DLT); 2 – if the DLT is reversible and 3 – if the DLT can be monitored clinically. Understanding mechanisms of toxicity helps extrapolate the relevance of the preclinical data to human. Taken together, preclinical findings are used to develop a margin of safety (MOS) approach based on the difference between the efficacy dose and the highest dose that does not cause toxicity, the no observed adverse effect level (NOAEL; [14]).

More often than not, preclinical data predict human risk [2]. Thus, it follows that increasing the preclinical MOS should improve clinical safety. This premise should hold even if people are more sensitive because a wider MOS makes it less probable that drug exposure required for efficacy will drift into a toxic range. If hepatic, cardiac and neurological toxicity contribute disproportionately to human toxicity, and if preclinical studies predict many of these outcomes [2], it also follows that improving the preclinical MOS for DLTs involving these three organs would result in fewer clinical adverse drug reactions. Thus, it is important to both *expand the view* to new preclinical approaches and *narrow the focus* on the selection of molecules with a wider MOS against the most problematic DLTs.

Screening strategies

Eliminating problem molecules early is the simplest way to avoid toxicity, but how and when should a screening strategy be implemented? By the time programs reach lead optimization (Fig. 2), chemical diversity in a structure activity relationship (SAR) narrows quickly to structures that are potent, bioavailable and effective in preclinical models. Once molecules are advanced into short duration *in vivo* toxicity testing, options to change the core chemical scaffold are limited because extensive optimization has already occurred. An optimal screening strategy is the one with higher throughput, precedes *in vivo* testing and can be deployed while the SAR affords ample degrees of freedom in selecting safer chemical structures. Early safety predictions in the absence of *in vivo* data can be met with skepticism, because the target organs and MOS are unknown [13]. To derive maximum benefit, however, predictive assays must be deployed before progressing to *in vivo* testing.

In designing a screening strategy, it is important to consider not only the target organ, but also broader classifications of toxicity. Toxicity can be on-target or off-target; the former may be due to target-based pharmacology, such as excessive bleeding with anticoagulants, and the latter to the chemical structure, for example acetaminophen-derived reactive metabolites. Toxicity can also be classified as cell autonomous or cell nonautonomous based on whether or not toxicity is due to a primary effect on the target cell or secondary to events in another cell or organ, respectively. For example, cardiac risk may be due to direct binding of drugs to the *hERG* channel on cardiomyocytes [12], or secondary to a centrally mediated increase in blood pressure. Cell autonomous toxicity can be addressed using cell-based screens, if the pathobiology can be replicated in a cell-based model. Screening for noncell autonomous toxicity requires complex and integrated biological system containing the relevant cell types.

Cell-based screening and cell autonomous toxicity

The higher throughput and low compound requirement for *in vitro* cell-based methods makes them attractive for screening and ranking candidate compounds [15,16]. All companies use some type of cell-based safety screens (e.g. chromosomal aberration testing for clastogens), but applications and strategies may differ. For cytotoxicity screening, there are two basic approaches, a universal screening approach using one or a few cell lines, or a target-organ-based approach, using cells with more specialized functions.

The major advantage of a universal cell-based screening strategy is that it can be used quickly to evaluate a high volume of compounds, because they generally use a single cell line with a high degree of automation [17]. There may be few preselection criteria with compounds gated into the assay directly from, for example, a compound solubility screen used to verify sufficient solubility to permit *in vitro* testing. The attraction of high throughput may be offset by less than favorable *in vitro* to *in vivo* correlations. Inaccurate prediction of *in vivo* outcome can lead to misapplication and/or a loss of confidence. The fact that the assay depends on a single cell type can also contribute to an increased error rate.

Another approach is to use primary cell culture models or more complex *in vitro* models, such as organ slices. Primary cell culture systems can be specifically designed to recapitulate the target

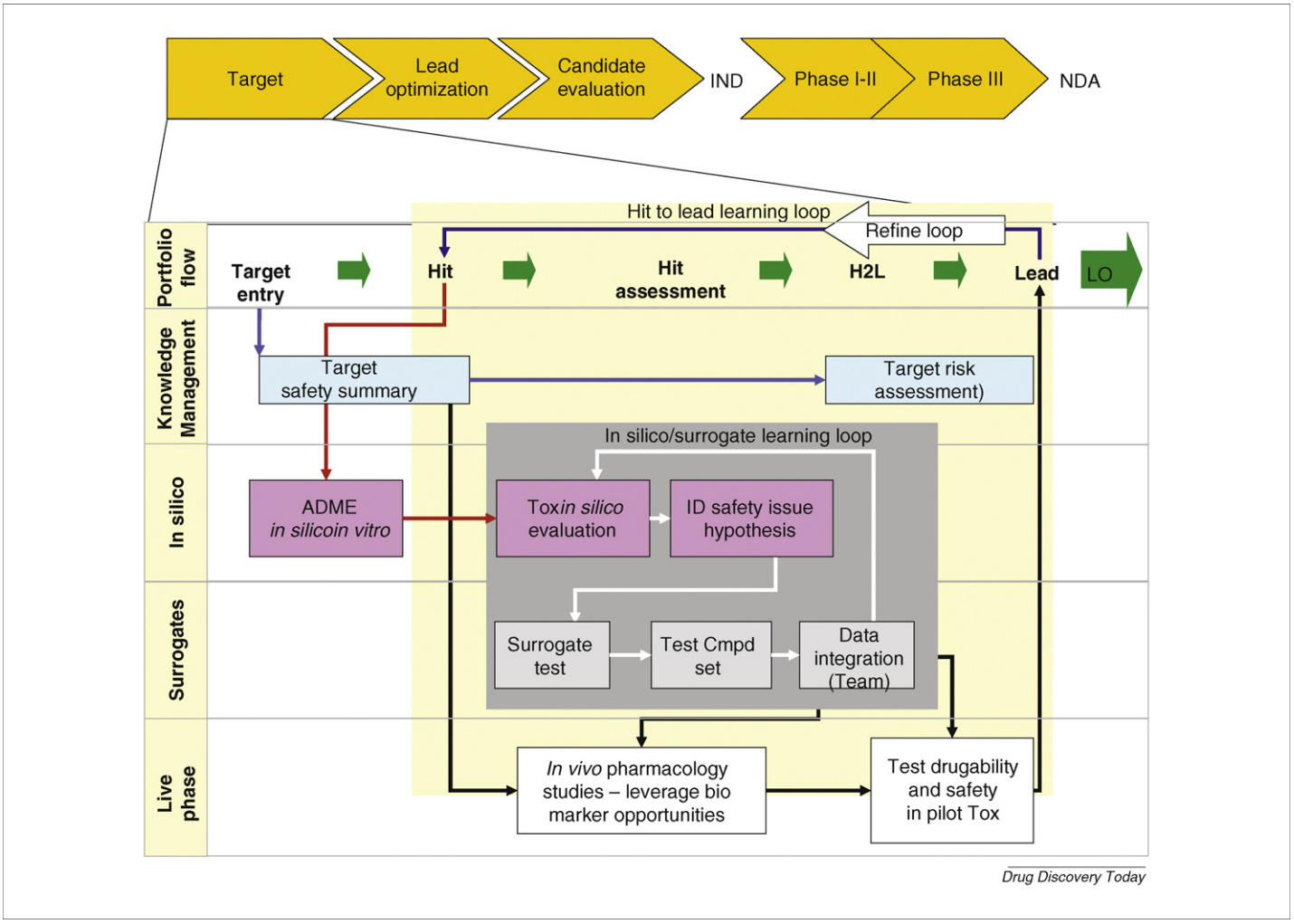

FIGURE 2

Opportunities for integrating surrogates screens into early risk assessment. The figure shows several lanes of activity (left), including surrogate assays, underneath an outline of the drug development process (top). The connectors within the work activities allow for a flexible approach that depends on the initial target assessment, *in silico* results from hits and surrogate assay data. Two learning loops are described, an inner loop through which *in silico* data are validated with surrogate results (gray box) and an outer loop for validation of the *in silico* and surrogates predictions in pilot *in vivo* rodent toxicity studies (yellow box). The focus is from target entry to lead declaration. Additional details on safety assessment strategies across the entire drug development pipeline can be found in available references [3,13,28,42]. IND, investigational new drug application; NDA, new drug application.

organ biology and are typically characterized against the intact organ to determine the level of functionality [15,18]. This potential for organ-specific responses is attractive if the strategy is to focus on areas of high attrition, such as liver and heart toxicity. Primary cells, however, may be more sensitive to handling and less amenable to automation and higher throughput. Nonetheless, these functionally defined models offer a solid foundation on which to build a successful cell-based screening strategy and can be expanded to technologies, such as high content imaging, genomics and proteomics [19,20].

Regardless of the model, the goal is to advance safer compounds with a wider MOS. Our approach is to make scaffold-based decisions early in discovery when chemistry remains structurally diverse. The objective is to steer the SAR toward a risk profile from which compounds can be advanced into *in vivo* safety assessment. It is crucial to recognize the limitations of cell models and incorporate that knowledge into decisions. No individual cell-based screening strategy is going to cover all situations; thus, a well-defined strategy focused on high value target organs generating sound biological data that is appropriately positioned in the drug development paradigm will offer the best opportunity to improve safety.

Surrogate animals and cell nonautonomous toxicity

A review of surrogate animal models is beyond our scope, but these models offer advantages in ease of genetic manipulation and throughput over traditional mammalian models. The use of animal surrogates for cell nonautonomous safety issues is gaining acceptance, but literature is scarce. Here again, a fit-for-purpose strategy, rather than a global screen, may be most effective. Excluding single cell organisms, *C. elegans* offer unparalleled ease of genetic manipulation; large genetic databases offer insight into development, mechanisms of cell death and offer candidate gene discovery capability [21]. Zebrafish offer the opportunity to collect observational data on intact organ systems and may be useful for safety pharmacology screening [22,23]. In the near term, surrogate

animal models can be used early in screening, but GLP studies in traditional preclinical species will support first studies in human for the foreseeable future.

Clinical focus: predicting, diagnosing and reacting to adverse drug reactions in human

As noted in the recent Institute of Medicine report [1], '...a drug's risk-benefit profile necessarily evolves over the drug's life cycle.' Postmarketing detection and monitoring for idiosyncratic drug reactions is beyond our scope, but a recent review by Uetrecht [5] addresses the challenges. Improved preclinical testing is unlikely to detect idiosyncratic adverse events unless new or genetically modified models are developed, a longer term approach. The clinical incidence for those adverse events predicted by animals [2,7,24] might, however, be decreased in the near term by improving the existing testing paradigm.

The case of drug-induced liver injury (DILI), along with the recent FDA draft guidance on DILI, and reference therein, helps frame the discussion [24]. The frequency of DILI is low during clinical testing, thus even large clinical trials are unlikely to estimate the true risk [24]. Hy's Law, the combination of elevated ALT/AST activity and total bilirubin concentration, is a strong predictor for DILI, but only 10% of Hy's Law patients will progress to irreversible liver injury. Even if no Hy's Law cases occur in 3000 subjects, one can only predict ($P < 0.05$) that the incidence of severe DILI is less than 1/10,000 [25]. With troglitazone, however, the incidence of acute liver failure was 2.4/10,000 and led to withdrawal [26]. Increasing the clinical population would increase the statistical power of the clinical design, but is impractical and would delay the development of novel treatments. With these constraints, two approaches can be considered to reduce the incidence of DILI; 1 – improve prediction accuracy and 2 – improve diagnosis and monitoring. Understanding mechanism and developing techniques to identify predisposed individuals are important to both approaches, but are beyond our scope and are addressed in recent reviews [5,12].

Improving preclinical predictions of human responses

Preclinical safety testing is, in the simplest sense, an assay designed to predict human safety. Performance for any assay can be evaluated on the basis of the rates for true (TP) and false positive (FP) versus true (TN) and false negative (FN) predictions. Preclinical assays are designed to protect patient safety by optimizing negative predictivity at the expense of higher FP rates. For example, only a small percentage of patients with long QT syndrome (LQTS) suffer a *torsades de pointe* (TdP), the associated fatal arrhythmia [12]. LQTS can occur if a drug binds with high affinity to the *hERG* voltage-gated potassium channel and delays repolarization [12]. Compounds likely to prolong QT are routinely excluded during *in vitro* preclinical testing. Because the frequency of *TdP* is low among all patients with LQTS, eliminating compounds that prolong the QT interval preclinically will have a high FP prediction rate for *TdP*. Compounds that cause *TdP* bind *h*ERG with high affinity and cause LQTS preclinically; thus, the assay has good negative predictivity (TN). Current guidelines for LQTS testing are biased toward optimizing negative predictions [27].

Now consider the assay parameters necessary to predict DILI. One study suggests that for all liver injury (regardless of the definition as nonidiosyncratic or idiosyncratic) noted during clinical trials, over half was predicted by preclinical data [2]. This, of course, does not include most of the idiosyncratic DILI events, because these typically are not detected in clinical trials. The vast majority of postmarketing cases are rare and potentially idiosyncratic DILI. By contrast, retrospective analysis suggests that when DILI is noted postmarketing, there are early signals in clinical trial data [7,24]. But only 10% of Hy's Law patients, those with altered liver function tests, progress to DILI [24]. Thus, among all compounds eliminated from development based only on altered liver function tests there may be a high number of FP predictions for DILI. Nonetheless, DILI for some marketed compounds was preceded by evidence of liver injury in clinical trials [24], thus, the FN rate may be low. In principle, even an assay optimized for TN predictions offers a way to advance compounds with a wider MOS and reduce the incidence of DILI in clinical trials. The trick is to implement the assay early enough that the FP prediction rate does not unduly prevent the advancement of new medicines for unmet medical needs. A strategy we employ for integrating surrogate screening for hepatotoxicity is shown in Fig. 2. The scheme includes *in silico* tools, which will not be addressed, but are integral to the strategy.

There is little direct evidence that an early preclinical intervention strategy will impact the incidence of DILI or other adverse drug reactions. However, an indirect case can be made from available data by working backward from clinical data to preclinical studies. Preclinical testing predicts a substantial percentage of human risk in clinical trials. When preclinical testing does predict human risk, studies of 30-day duration or less effectively identify the vast majority [2]. Thus, improving MOS in 30-day studies should reduce the incidence of preventable DILI, and other adverse reactions, in clinical trials. Going back one more step, shorter term 4–14 day *in vivo* studies often predict outcome in 30-day studies [13,28]. Finally, if hepatocytes *in vitro* predict liver safety in short duration *in vivo* studies, then early screening should allow safer compounds to be advanced. Our experience with predicting short-term *in vivo* outcomes with rat primary hepatocytes (RPHs) screening is summarized in Table 1. The data support the argument that a well-connected set of assays deployed from early *in vitro* intervention through the clinic offers an opportunity to improve safety by selecting better molecules. This strategy requires early intervention to offset the penalty imposed by the

TABLE 1

Performance of primary rat hepatocytes in predicting liver injury.

	Positive *in vivo*	Negative *in vivo*	Total	Accuracy
RPH predicted positive	13	7	20	65%
RPH predicted negative	2	17	19	90%
Total	15	24	39	
%Sensitivity; specificity	87%	71%		

The table shows the performance of rat primary hepatocytes (RPHs) in predicting >2X increases in ALT activity in serum in four day rat repeat dose studies during lead optimization. Row labels at the left (predicted positive or negative) are for RPHs and were based on cytolethality scores after being treated with compounds in culture. Positive and negative (>2X ALT increase) outcomes *in vivo* are identified at the top of each column and are based on data from each study. The data were accumulated from compounds progressing through lead optimization as part of the normal drug development process.

FP rate with chemical degrees of freedom available in the SAR. It is also important to emphasize again that this strategy is not designed to prevent idiosyncratic adverse drug reactions and will be most effective for cell autonomous toxicity.

Better biomarkers for prediction, detection and diagnosis

Biomarkers can be biochemical assays, antibody-based techniques, imaging modalities and genomic markers. Noninvasive biomarkers are most useful for clinical monitoring; preclinical applications offer added flexibility because tissues are routinely collected allowing more invasive approaches. Ideal biomarkers predict toxicity before injury occurs or in a time window when the injury is reversible. For example, increased ALT/AST activity in serum occurs early when hepatocytes are injured or die, but adaptive responses can prevent loss of liver function. Alternatively, injury can progress to loss of function, as reflected by increased ALT/AST activity plus total-bilirubin concentration, during which time the risk of liver failure increases markedly [25]. Biomarkers that predicted which patient would progress from increased ALT/AST activity through increased total-bilirubin concentration to liver failure would significantly improve the ability to monitor patients at risk for DILI.

Biomarkers for cardiac risk include classical biomarkers, such as creatine kinase and AST as well as newer biomarkers, such as cardiac troponins and the natiuretic peptides BNP and NT-pro-BNP [29,30]. Cardiac troponin I is released from injured cardiomyocytes upon injury or death. BNP and NT-pro-BNP increase when pro-BNP expression and processing increases owing to increased ventricular stretch. The difficulty lies in knowing when increased troponin I concentration or BNP is associated with cardiac adaptation, for example to exercise, or predicts risk of cardiotoxicity. New biomarkers may add granularity to cardiac risk assessment and allow better surveillance methods to avoid injury rather than diagnose that which has occurred.

Regardless of whether a biomarker is a metabolite, an RNA transcript, or a behavioral observation, the challenge is to improve monitoring so that irreversible organ damage is avoided. Development of novel biomarkers and biomarker panels in the near term can have an immediate impact both on patient safety and drug pipeline productivity.

The future: personalized medicine, pharmacogenomics and systems biology

We outlined a framework for approaching early drug safety assessment. We suggested a narrow focus on a few target organs and an expanded view to understanding the impact of early intervention on adverse drug reactions. This near term, but not short sighted, focus will not address truly idiosyncratic toxicities not predicted by preclinical or clinical testing. A better understanding of mechanisms is necessary to frame the problem of predicting idiosyncratic toxicity as testable hypotheses [5]. Nontechnical issues also contribute to the overall burden of drug toxicity. For example, adverse drug reactions are under reported [31] suggesting that improvements in patient monitoring and detection are also needed [6,32]. Recent drug withdrawals enhance the perception that newer drugs have a higher incidence of adverse reactions [33], but over-the-counter medications, for example acetaminophen, and commonly used prescription drugs, for example warfarins, statins and antiretrovirals, for which there are few or no substitutes, contribute significantly to the overall incidence [6,12,31,34].

Advances in pharmacogenomics/pharmacogenetics and systems biology may reveal new mechanisms of disease and drug susceptibility [12,35,36]. Genetic differences in metabolism and disposition have already been identified for marketed drugs [12,34,36]. Sensitivity to some adverse reactions maps to MHC genotype supporting an immune-mediated mechanism [5,37]. Newer genetic tests allow dosing regimens for drugs with known susceptibility factors to be fine tuned [38,39] and candidate gene approaches may clarify the genetic basis of mono- and polygenic susceptibility traits [38,40]. The mouse phenome project and the Pharmacogenomic Research Network are linking genotype to phenotypes and vice versa in rodent and human, respectively [36,41]. Much work is required to realize the value of pharmacogenetics in predicting and detecting adverse drug reactions [36], but progress is being made.

Summary and conclusions

The easiest way to improve drug safety is to avoid issues by selecting better molecules; easy to say and difficult to do. Chemical diversity narrows quickly leaving a short time window of opportunity relative to the time necessary to deliver a safer drug to patients. Expanding the view to making better decisions early while narrowing the focus to addressing the most important safety issues is crucial to improving drug safety. The imperative is clear, but progress requires integrating short-term gains within a longer term focus. The strategies addressed in this discussion are one piece of the puzzle. Points to consider are summarized in Box 1.

BOX 1

Summary of points to consider for improving drug safety

- Expand application of predictive approaches early in discovery.
- Narrow the focus on the most important target organs.
- Identify on- versus off-target and cell-autonomous versus cell nonautonomous screening strategies.
  - Cell culture models
  - Surrogate animal models
- Make scaffold-based decisions early using surrogate data.
- Improve preclinical margins of safety in studies of 30-day duration or less.
- TNs must be balanced against the FP rate.
  - Preclinical tests are often biased toward TN predictions.
- Develop biomarkers to address gaps in preclinical and clinical monitoring.
- Apply pharmacogenomic screens to avoid drug-sensitive populations.
- Differentiate strategies for nonidiosyncratic and idiosyncratic drug toxicity.
- Over-the-counter medications and prescription drugs without good substitutes contribute significantly to adverse drug reactions.

Acknowledgements

The authors thank Michael Dorato, Derek Leishman, Carl McMillian, Marija Popovic, Arie Regev, Eric Schultze, Anja Stauber and Craig Thomas for their critical reading of the manuscript and thoughtful commentary. We also thank Craig Zwickl and Brian Mattioni for their assistance with the analysis in Table 1.

References

1 Committee on the Assessment of the US Drug Safety System, *et al.* (2007) *The Future of Drug Safety: Promoting and Protecting the Health of the Public*. National Academy Press
2 Olson, H. *et al.* (2000) Concordance of the toxicity of pharmaceuticals in humans and in animals. *Regul. Toxicol. Pharmacol.* 32, 56–67
3 Stevens, J.L. (2006) Future of toxicology – mechanisms of toxicity and drug safety: where do we go from here? *Chem. Res. Toxicol.* 19, 1393–1401
4 Schmid, E.F. and Smith, D.A. (2004) Is pharmaceutical R&D just a game of chance or can strategy make a difference? *Drug Discov. Today* 9, 18–26
5 Uetrecht, J. (2008) Idiosyncratic drug reactions: past, present, and future. *Chem. Res. Toxicol.* 21, 84–92
6 Pirmohamed, M. *et al.* (2004) Adverse drug reactions as cause of admission to hospital: prospective analysis of 18,820 patients. *BMJ* 329, 15–19
7 Alden, C.L. and Lin, J.S.P. (2003) Predictive toxicology technology for avoiding idiosyncratic liver injury. *Preclinica* 27–35 May/June
8 Shah, R.R. (2006) Can pharmacogenetics help rescue drugs withdrawn from the market? *Pharmacogenomics* 7, 889–908
9 Fung, M. *et al.* (2001) Evaluation of the characteristics of safety withdrawal of prescription drugs from worldwide pharmaceutical markets – 1960 to 1999. *Drug Inform. J.* 35, 293–317
10 Lasser, K.E. *et al.* (2002) Timing of new black box warnings and withdrawals for prescription medications. *JAMA* 287, 2215–2220
11 CDER Report to the Nation: Improving Public Health through Human Drugs. http://www.fda.gov/CDER/about/default.html
12 Wilke, R.A. *et al.* (2007) Identifying genetic risk factors for serious adverse drug reactions: current progress and challenges. *Nat. Rev. Drug Discov.* 6, 904–916
13 Kramer, J.A. *et al.* (2007) The application of discovery toxicology and pathology towards the design of safer pharmaceutical lead candidates. *Nat. Rev. Drug Discov.* 6, 636–649
14 Estimating Maximum Safe Starting does in Initial Clinical Trials for Therapeutics in Adult Healthy Volunteers. Center for Drug Evaluation and Research http://www.fda.gov/cder/Guidance/5541fnl.html. 2005
15 Guengerich, F.P. and MacDonald, J.S. (2007) Applying mechanisms of chemical toxicity to predict drug safety. *Chem. Res. Toxicol.* 20, 344–369
16 Eisenbrand, G. *et al.* (2002) Methods of in vitro toxicology. *Food Chem. Toxicol.* 40, 193–236
17 Evans, S.M. *et al.* (2001) Development of a high throughput in vitro toxicity screen predictive of high acute in vivo toxic potential. *Toxicol. In Vitro* 15, 579–584
18 Johnson, T.E. *et al.* (2004) Statins induce apoptosis in rat and human myotube cultures by inhibiting protein geranylgeranylation but not ubiquinone. *Toxicol. Appl. Pharmacol.* 200, 237–250
19 Gough, A.H. and Johnston, P.A. (2007) Requirements, features, and performance of high content screening platforms. *Methods Mol. Biol.* 356, 41–61
20 Ekins, S. *et al.* (2005) Techniques: application of systems biology to absorption, distribution, metabolism, excretion and toxicity. *Trends Pharmacol. Sci.* 26, 202–209
21 Segalat, L. (2007) Invertebrate animal models of diseases as screening tools in drug discovery. *ACS Chem. Biol.* 2, 231–236
22 Rubinstein, A.L. (2006) Zebrafish assays for drug toxicity screening. *Expert. Opin. Drug Metab. Toxicol.* 2, 231–240
23 Zon, L.I. and Peterson, R.T. (2005) In vivo drug discovery in the zebrafish. *Nat. Rev. Drug Discov.* 4, 35–44
24 Drug-Induced Liver Injury: Premarketing Clinical Evaluation. Center for Drug Evaluation and Research http://www.fda.gov/cder/guidance/7507dft.html. 2007
25 Temple, R. (2006) Hy's law: predicting serious hepatotoxicity. *Pharmacoepidemiol. Drug Saf.* 15, 241–243
26 Graham, D.J. *et al.* (2003) Incidence of idiopathic acute liver failure and hospitalized liver injury in patients treated with troglitazone. *Am. J. Gastroenterol.* 98, 175–179
27 Center for Drug Evaluation and Review. The Non-clinical Evaluation of the Potential for Delayed Ventricular Repolarization (QT Interval Prolongation) by Human Pharmaceuticals. http://www.fda.gov/cber/gdlns/ichqt.html. 2005
28 Sasseville, V.G. *et al.* (2004) Testing paradigm for prediction of development-limiting barriers and human drug toxicity. *Chem. Biol. Interact.* 150, 9–25
29 Hall, C. (2004) Essential biochemistry and physiology of (NT-pro)BNP. *Eur. J. Heart Fail.* 6, 257–260
30 Maisel, A.S. *et al.* (2006) Cardiac biomarkers: a contemporary status report. *Nat. Clin. Pract. Cardiovasc. Med.* 3, 24–34
31 Hazell, L. and Shakir, S.A. (2006) Under-reporting of adverse drug reactions: a systematic review. *Drug Saf.* 29, 385–396
32 Hauben, M. and Reich, L. (2005) Potential utility of data-mining algorithms for early detection of potentially fatal/disabling adverse drug reactions: a retrospective evaluation. *J. Clin. Pharmacol.* 45, 378–384
33 Moore, T.J. *et al.* (2007) Serious adverse drug events reported to the Food and Drug Administration, 1998–2005. *Arch. Intern. Med.* 167, 1752–1759
34 Weinshilboum, R. and Wang, L. (2004) Pharmacogenomics: bench to bedside. *Nat. Rev. Drug Discov.* 3, 739–748
35 Weinshilboum, R.M. and Wang, L. (2006) Pharmacogenetics and pharmacogenomics: development, science, and translation. *Annu. Rev. Genomics Hum. Genet.* 7, 223–245
36 Giacomini, K.M. *et al.* (2007) The pharmacogenetics research network: from SNP discovery to clinical drug response. *Clin. Pharmacol. Ther.* 81, 328–345
37 Fox, J. *et al.* (2006) The clinical implications of antiretroviral pharmacogenomics. *Pharmacogenomics* 7, 587–596
38 Guo, Y. *et al.* (2006) In silico pharmacogenetics of warfarin metabolism. *Nat. Biotechnol.* 24, 531–536
39 Wu, A.H. *et al.* (2008) Dosing algorithm for warfarin using CYP2C9 and VKORC1 genotyping from a multi-ethnic population: comparison with other equations. *Pharmacogenomics* 9, 169–178
40 Gatti, D. *et al.* (2007) Genome-level analysis of genetic regulation of liver gene expression networks. *Hepatology* 46, 548–557
41 Bogue, M.A. *et al.* (2007) Mouse Phenome Database (MPD). *Nucleic Acids Res.* 35, D643–D649
42 Mayne, J.T. *et al.* (2006) Informed toxicity assessment in drug discovery: systems-based toxicology. *Curr. Opin. Drug Discov. Dev.* 9, 75–83