Clarivate Cortellis™

Special report

Best practices in toxicology

Current challenges and novel approaches for enhancing drug safety

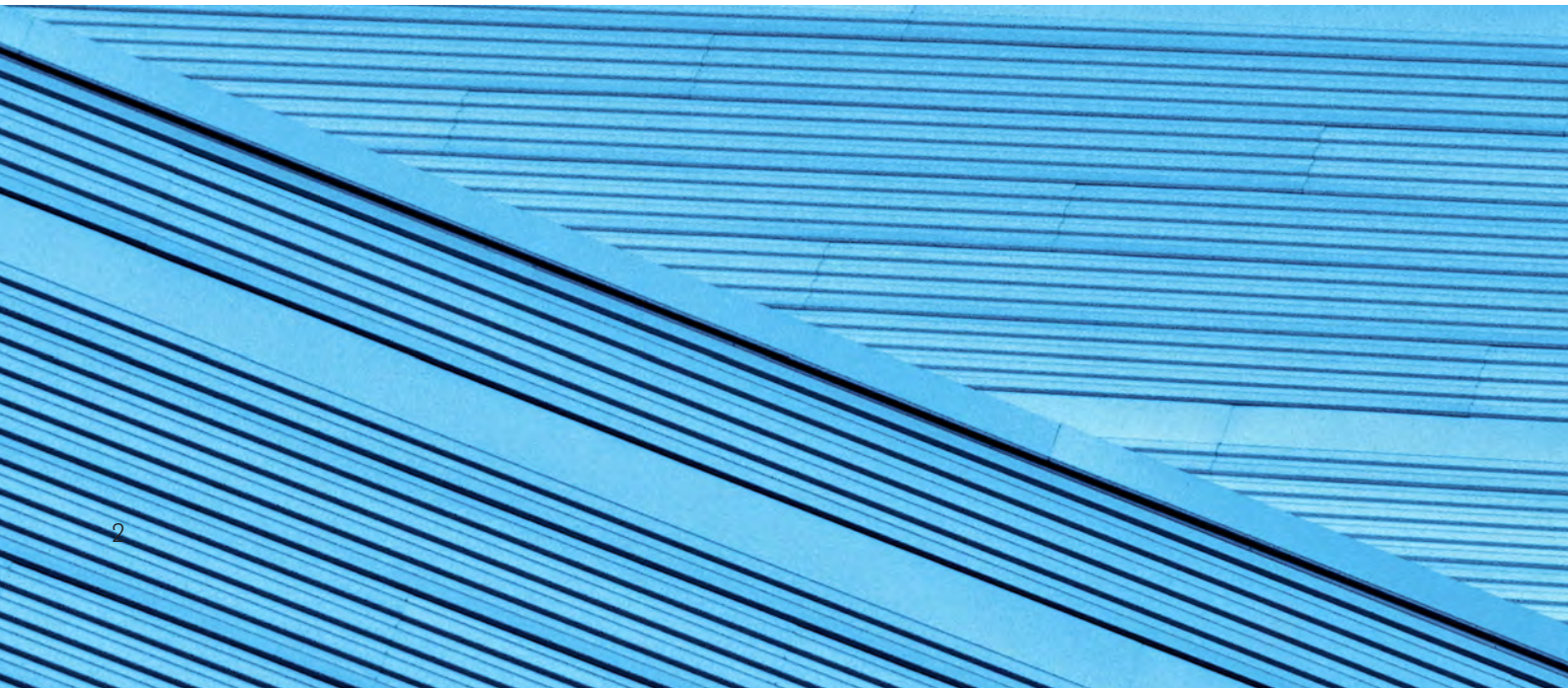
Toxicity at the forefront

The global toxicity testing market is projected to reach $23.55 billion in the next 2 years.[1]

A wide range of industries producing food, personal care products and agricultural commodities rely on toxicity testing to ensure the safety of these products for human use and/or consumption. The biotechnology and pharmaceutical markets might represent a couple of the biggest drivers of this market. It is estimated that for every eight active substances entering Phase I clinical trials, only one product is approved and marketed.[2] The low success rate for clinical trials is largely driven by two factors: poor efficacy in humans and adverse events leading to the discontinuation of a drug development program.[3] Unanticipated preclinical toxicity and adverse events in the clinical stage are leading causes of drug attrition, directly contributing to the rapidly increasing costs of drug development.

A variety of factors contribute to adverse outcomes for a given therapy. Largely, these are determined by the inherent properties of the therapy itself, such as toxicity related to on and off-target effects, metabolism and pharmacokinetics of a compound. However, there is also a growing appreciation for the nuanced role of genetics in determining the extent of an adverse outcome for a given individual. Collectively, these findings can be used to improve the overall success rates of a drug development program. In the sections below, we discuss how each step in the drug discovery phase is crucial to improving the success rates and some of the best practices that could be incorporated into the workflow.

Addressing off-target effects and considering the role of genetics in adverse outcomes has the potential to improve overall success rates of a drug development program.



Starting early: identifying risks at the discovery stage

For every drug that is approved by a regulatory agency, 5,000 to 10,000 chemicals are screened in the early discovery stage, of which only a few hundred are pursued for preclinical development.[3]

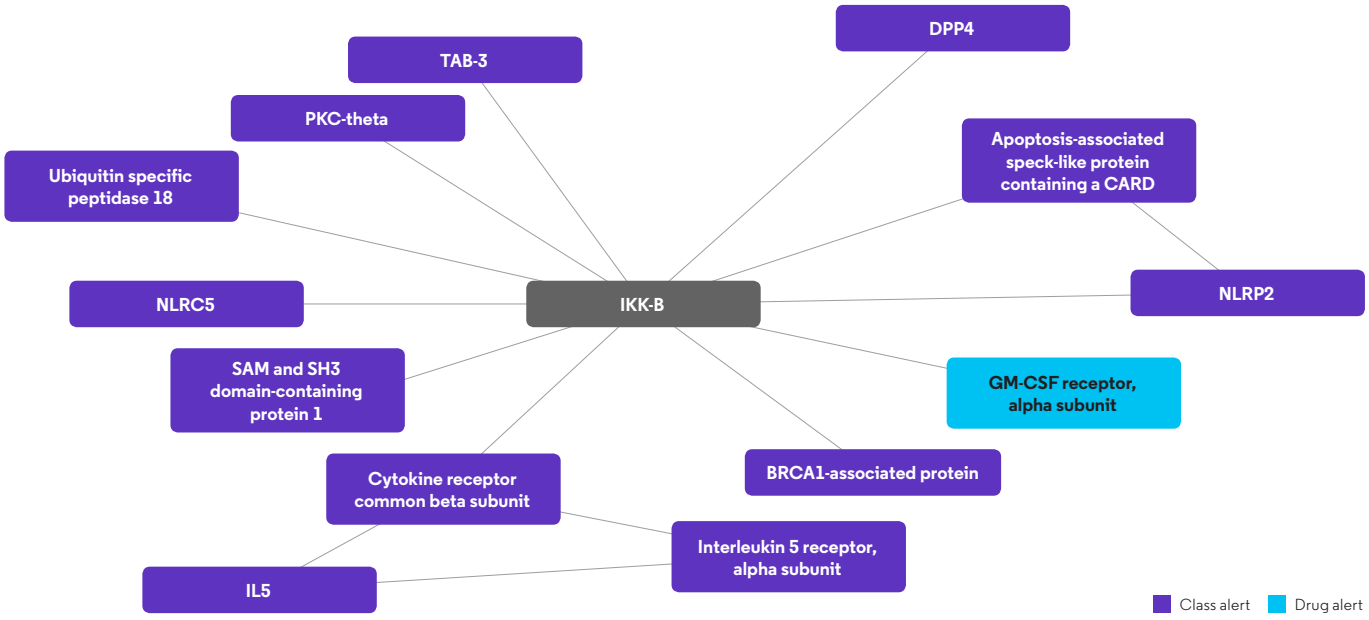
With global R&D expenditure projected to reach $92 billion by 2022, investors are eager to ensure that their money is well-spent by advancing the most robust and least risky candidate to the clinic.[2] It is therefore necessary to critically evaluate all aspects of this stage that could potentially impact the overall cost and time to identify a lead candidate. From a safety perspective, some of the important considerations include target liability; analysis; drug design, lead identification; and drug repurposing.

Target liability

The foundation for early discovery research is almost always published literature. Thus, the focus at this stage is to replicate the validity of existing evidence and identify druggable targets and/or disease areas for drug development. While establishing proof of concept can be a challenging task in itself, potential safety liabilities for a given target need to be assessed in addition to determining whether pursuit of a novel target, in an effort to be first-in-class, could result in wasted time, money and resources if the drug is associated with adverse side effects.

An inherent risk with novel targets is that not much is typically known about associated toxicity, aside from pre-existing data (if any) from gene silencing or knockout studies. At a preliminary level, using a network-based approach to understand the target neighborhood could offer unique insights into potential mechanistic pathways for adverse outcomes and mitigation plans. The strategy then is not to necessarily abandon the target but to anticipate and be prepared to address issues arising from preclinical toxicity or adverse events at the clinical stage. As shown in Figure 1, neighborhood analysis of an emerging target, IKKB, shows its direct interaction with many targets associated with class alerts. Considering this information during therapeutic development can help mitigate unintended risks that result from a positive or negative impact to these neighborhood relationships.

Figure 1: Neighborhood analysis of the target IKK-B.



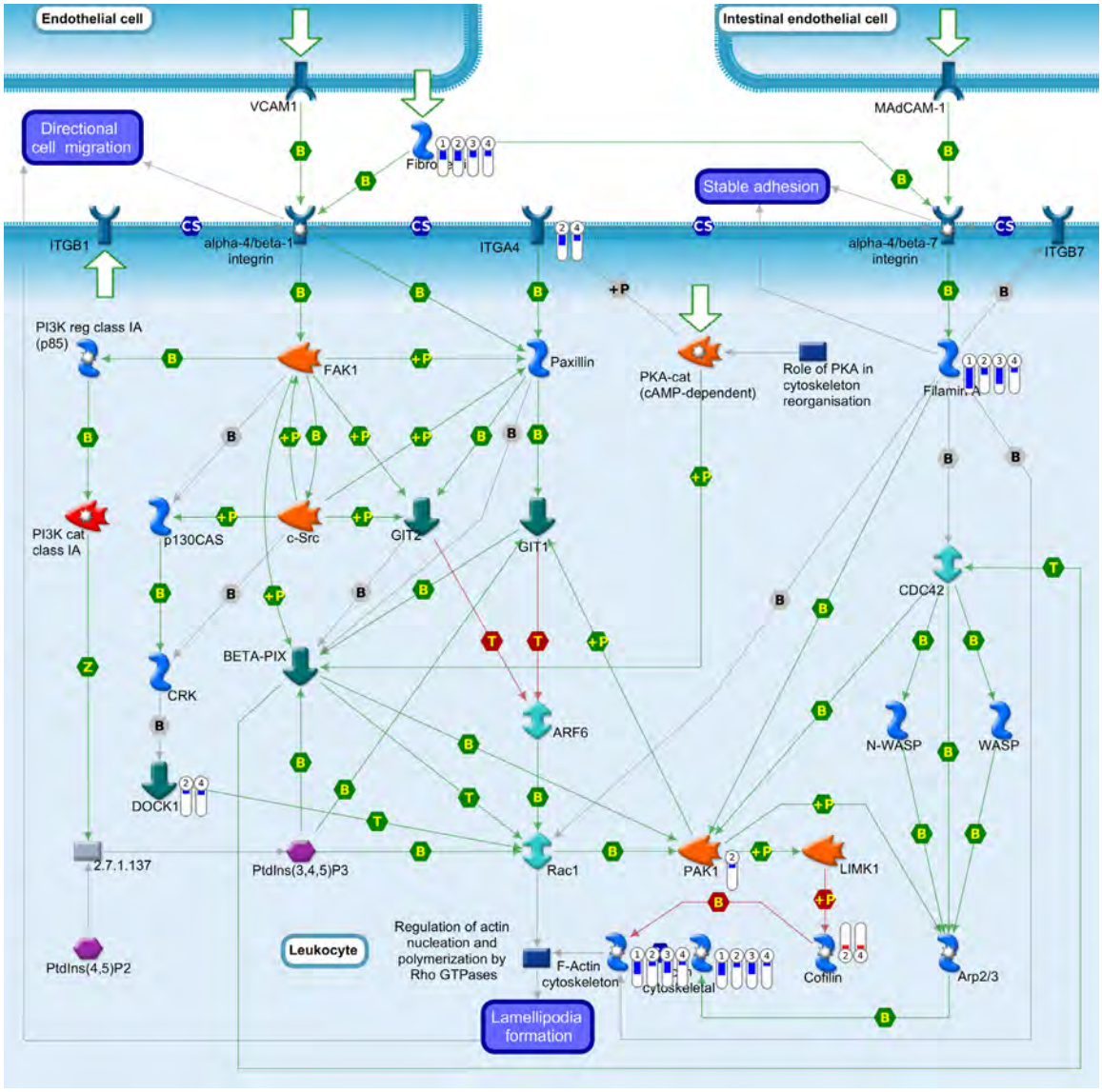
Targets in purple have known class alerts while the target in blue has a safety alert associated with a drug against it. Known class alerts are used when the target family is linked to adverse events demonstrated in preclinical knockdown/knockout models or have known human gene-disease association.

Source: Bioinfogate OFF-X

Integrated approach to data analysis

A more holistic approach to understanding preclinical toxicity is combining multi-OMICs datasets (genomics, transcriptomic, metabolomic and proteomic) to identify toxicity-specific biomarkers and pathways impacted by treatment with a candidate drug. Rather than assessing neighborhoods around a single target or biomarker of interest, this approach helps understand overall pathways and processes perturbed by a drug (Figure 2).[4,5] Assessing treatment outcomes at multiple molecular levels not only offers cross-validation of the experimental data but also helps identify key players previously unknown to the toxicity mechanism. Integrative use of multi-OMICs data was previously shown to identify mechanisms and biomarkers for liver injury associated with drugs such as cyclosporin S and bosentan.[6]

Figure 2: Integrative data analysis to identify biological pathways with greatest impact from a drug treatment.



This top ranking pathway map from the combined analysis of proteomics (experiments 1 and 3) and gene expression (experiments 2 and 4) data obtained at day 7 (experiments 1 and 2) and day 14 (experiments 3 and 4) of a treatment of 450nM doxorubicin to stem cell-derived cardiomyocytes indicates a perturbation in α-4 integrins-mediated cell adhesion and migration. Upwards or downwards pointing thermometers in red and blue indicate upregulation or downregulation, respectively.

Source: MetaCore, a Cortellis™ solution

Drug design and lead identification

Identifying lead candidates is perhaps the most important step of the drug discovery process. One of these contenders could ultimately be the next blockbuster drug or conversely contribute to the grim drug failure rate statistics. The main objectives of lead identification are to identify drugs of the highest potency that are well-tolerated in the experimental system. There are several strategies that can be deployed to identify leads to progress for preclinical development, including a target-centric approach and high-throughput phenotypic screens.

Target-centric approach

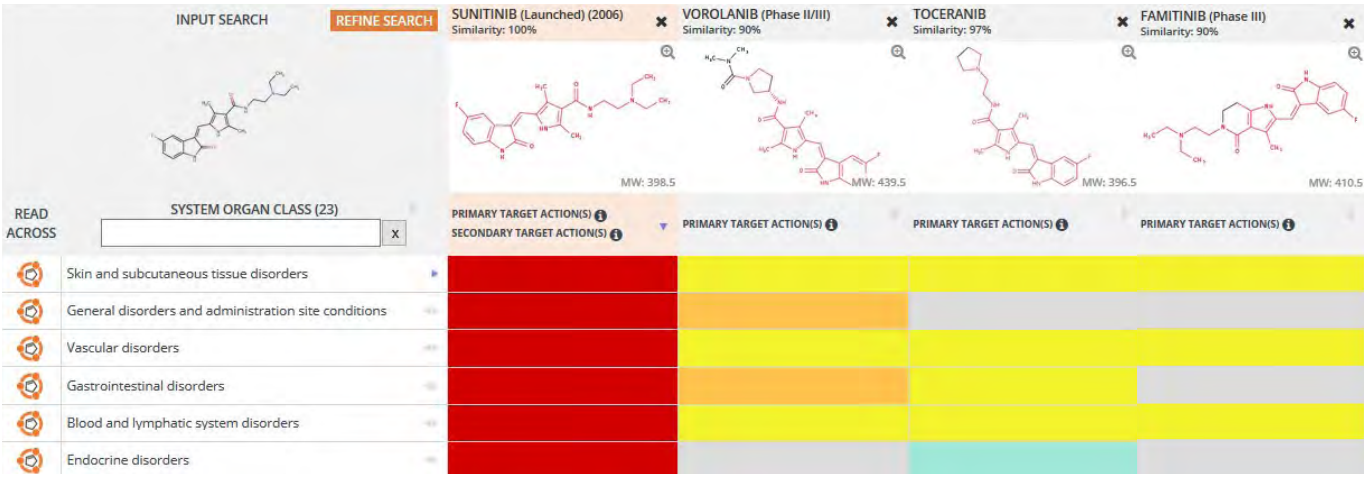
The target-centric approach involves choosing a well-studied target of interest for therapeutic advancement. This is the classical approach to drug discovery and, until recently, was the most popular viewpoint in the industry. By combining the information available from the 3-dimensional structure of a target (e.g., from X-ray crystallography, NMR spectroscopy or electron microscopy) and computer-aided drug design, compounds modulating the activity of a given target can be virtually created and tested experimentally. A similar concept can then also be used to identify the "off-target" effects. For instance, protein-ligand docking analysis of the HIV protease inhibitor nelfinavir revealed its ability to inhibit a variety of kinases involved in cancer, thereby providing a hypothesis for its observed anti-cancer effects.[7] This type of analysis would also be beneficial in predicting the toxicity of new chemical structures based on their predicted off-target interactions, especially when the adverse outcomes associated with the off-targets are already well-known or previously established.[8]

High-throughput phenotypic screens

Advances in areas from robotics and microfluidics to chemical and stem cell biology have collectively led to the rise and success of high-throughput screens (HTS). Screening millions of compounds on a wide array of substrates offers a faster, cheaper method of identifying a collection of therapeutic candidates. HTS was credited with the identification of antimicrobials that are new molecular entities (NMEs) with a novel mechanism of action.[9] The HTS approach has also simplified running large-scale toxicity screens for compound libraries and are routinely employed in the workflows (e.g., inhibition of the alpha subunit of the potassium channel or hERG). As the step prior to the HTS process, medicinal chemists are increasingly relying on in silico tools for virtual library screening using quantitative structure activity relationship (QSAR) and quantitative structure in vitro in vivo relationship (QSIIR) to make predictions on therapeutics properties as well as absorption, distribution, metabolism, excretion and toxicity (ADMETox).

Figure 4 (see next page) shows QSAR protein binding and toxicity predictions for the compound doxorubicin hydrochloride. These predictions help triage compounds entering HTS. While these tools themselves are a valuable resource, care should be given to also incorporate workflows for due diligence on pharmacology data, off-target effects and success and failure rates of similar chemical entitites - as seen in Figure 3.

Figure 3: Structure similarity and comparative safety profiles for the drugs sunitinib, vorolanib, toceranib and famitinib.



Source: Bioinfogate OFF-X

Figure 4: QSAR predictions for doxorubicin hydrochloride.

Property	Model description	Value	(TP)
Pgp-inh, pIC50	Human P-glycoprotein transporter inhibition, pIC50 (uM). Cutoff is -1.7. The higher the value, the higher the inhibition activity. Model description: N=274, R2=0.85, RMSE=0.4.	-0.81	(59.88)
Pgp-sub, prob	Potential to be a substrate of human P-glycoprotein transporter, range from 0 to 1. Cutoff is 0.5. Values closer to 1 indicate potential ligands. Reference: Penzotti, Lamb, et al., 2002 (PMID: 11960484). Model description: N=192, R2=0.65, RMSE=0.3.	0.93	(99.63)

A. Protein binding QSAR models: values greater than the listed cutoff indicate the high likelihood of the compound inhibiting and being a potential substrate for the P-glycoprotein. Values in parenthesis (TP) refer to the Tanimoto prioritization.

Property	Model description	Value	(TP)
Cardiotoxicity	Potential for inducing cardiotoxicity. Cutoff is 0.5. Values higher than 0.5 indicate potentially toxic compounds. Training set consists of chemicals and drugs causing cardiotoxicity in vivo. Model organisms: mouse, rat, human. Model description: Training set N=143, Test set N=30, Sensitivity= 0.80, Specificity=1.00, Accuracy=0.90, MCC=0.82. Reference: Clarivate Analytics, click here to view training set.	0.94	(99.63)
Cytotoxicity model, -log GI50 (M)	Growth inhibition of MCF7 cell line (human caucasian breast adenocarcinoma), pGI50. Cutoff is 6. Values from 6 to 8 correspond to a toxic metabolite, values less than 6 are preferable, values less than 3 are more preferable and less toxic. Reference: DTP/NCI. Model description: N=1474, R2=0.9, RMSE=0.05.	6.67	(99.63)
Epididymis toxicity	Potential for inducing epididymis toxicity. Training set consists of chemicals and drugs causing epididymis toxicity in vivo. Model organisms: mouse, rat, human. Cutoff is 0.5. Values higher than 0.5 indicate potentially toxic compounds. Model description: Training set N=252, Test set N=42, Sensitivity= 0.90, Specificity=0.86, Accuracy=0.88, MCC=0.76.	0.8	(99.63)

B. Toxicity prediction: values in red, greater than the cutoff, indicate the high likelihood of the compound exhibiting cardiotoxicity, cytotoxicity and epididymis toxicity. Values in parenthesis (TP) refer to the Tanimoto prioritization.

Source: MetaCore, a Cortellis solution

Drug repurposing

As drug development costs soar, finding a new use for an old drug is widely being embraced by drug developers. For drugs demonstrated to be safe and well-tolerated in clinical trials but failed to show efficacy, exploring potency in other disease areas as monotherapy or a part of combination therapy is an appealing alternative to the time and money invested to discover new drug candidates. This practice is already in place by pharmaceutical companies looking to extend revenues and patent life of their invested products.

Correspondingly, drugs with known safety issues can be repurposed with minimal or no side effects for an application different from their original intended use. While successful repurposing of a drug requires access to chemical libraries and drug databases,[10] supporting information on assays, animal models, pharmacology, clinical trials and safety data can greatly save time and improve the efficacy of the repurposing process. Figure 5 on the following page shows a list of the top 10 potential targets for non-alcoholic steatohepatitis (NASH) that have drugs available or being explored for other indications.

Pharmaceutical companies are looking to repurpose existing drugs to extend revenues and patent life of their products.

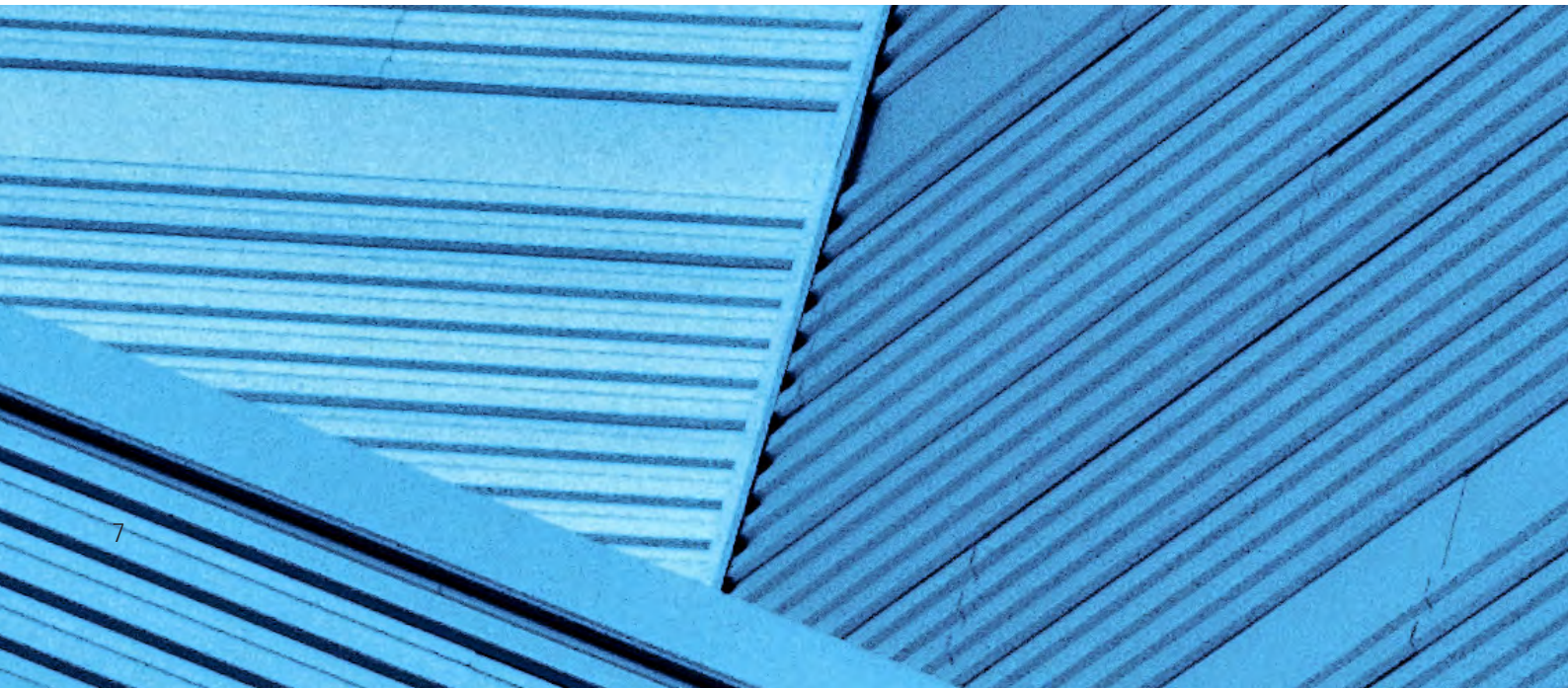


Figure 5: Top nine targets ranked by condition novelty for non-alcoholic steatohepatitis (NASH).

Rank	Target main name	Gene symbol	Drugs		Experimental pharmacology	Animal models		Biomarker uses		Genetic evidence	
			filtered	total		filtered	total	filtered	total	filtered	total
1	**Interleukin-1 beta**	IL1B	0	768	623	0	119	26	4973	1	614
2	**Estrogen receptor**	ESR1	0	757	2462	0	87	0	2055	2	781
3	**Signal transducer and activator of transcription 3**	STAT3	0	490	1233	0	105	8	1995	1	362
4	**Apolipoprotein E**	APOE	0	20	26	8	1264	5	973	1	891
5	**Advanced glycosylation end product-specific receptor**	AGER	0	102	72	0	77	5	885	2	179
6	**Microsomal triglyceride transfer protein large subunit**	MTTP	0	366	181	2	16	2	97	1	85
7	**1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase gamma-1**	PLCG1	0	10	111	0	1	0	280	1	65
8	**Alpha-1-antitrypsin**	SERPINA1	0	20	0	0	10	7	1054	1	104
9	**HLA class I histocompatibility antigen, A-1 alpha chain**	HLA-A	0	23	40	0	42	0	463	1	398

The table lists the targets and number of records with evidence supporting each heading in the columns. Values associated with the total number represent all disease areas including NASH.

Source: Cortellis Drug Discovery Intelligence

Understanding organ-specific toxicity while reducing the burden on animal testing

Animal testing increases the cost of drug development, and there are growing ethical concerns on the use of animals for biomedical research. In addition, underlying differences in biological and metabolic pathways between animal models and humans have often-times resulted in lack of translatability. These differences manifest as poor drug efficacy in humans and/or clinical stage adverse events, both of which significantly contribute towards drug attrition.

Implementing the 3Rs - Replace, Reduce and Refine - by replacing animal testing with in vitro and in silico approaches, reducing the need for animal testing and refining experimental design, can greatly diminish the burden of animal testing. In 2013, a ban on vertebrate animal testing for European cosmetic ingredients was imposed in accordance with the 7th amendment of the European Cosmetic Directive. A subsequent report from the transatlantic think tank on toxicology has laid out guidelines for alternative (animal-free) methods for toxicity testing.[11] One of the main highlights calls for the use of an integrative testing strategy, where one methodology is not deemed as the gold standard. Instead, the emphasis is to facilitate multiple in vitro testing methods to draw conclusions and to model and predict adverse outcomes with 100% accuracy using a system's biology approach.

Indeed, as described in the previous section, high-throughput screening has significantly impacted our ability to quickly test libraries of compounds at multiple dosing intervals. Moreover, recent advances in stem cell, gene editing and microfluidics technologies has made it easier to execute 2D, 3D and cell co-culturing techniques on a large scale. When combined with the right cell/tissue/organ assay, toxicity studies can be performed in large scale in vitro to recreate relevant tissue microenvironments. The approach of collecting a large number of datapoints would significantly improve the development of more accurate prediction models and provide a mechanistic understanding of pathways underlying adverse events (adverse outcome pathways [AoPs]). Such a mechanistic understanding would greatly benefit researchers to reliably make informed go or no-go decisions when developing NMEs or investigating the potential to repurpose drugs. Concurrently, the US Environmental Protection Agency's ToxCast and Tox21 programs aim to achieve a mechanistic understanding behind toxicities related to environmental and industrial chemicals, with the hope that these may serve as a tool to develop regulatory strategies in the future.[12]

The personalized medicine approach

Next-generation sequencing technology has advanced at an incredibly fast pace, leading to low costs in whole-genome sequencing and the ability to make multitudes of comparisons based on a broad range of factors. This has ushered in the era of toxicogenomics (and pharmacogenomics), which has provided researchers with large quantities of OMICs datasets to examine the mechanistic role of a drug or a toxicant in adverse events. One example of this would be assessing the influence of gene variants on the risk for adverse events. Identifying this relationship would help determine a patient segment that would benefit most from a given drug and the maximum tolerated dose for a given individual. A well-known example of the influence of genetic variability on drug-induced adverse events is the observed cardiotoxicity in women receiving doxorubicin chemotherapy, which is exacerbated when used in combination with Trastuzumab.[13,14]

Similarly, there is growing interest in understanding how epigenetics plays a role in determining susceptibility or increased risk for toxicity – also called pharmacoepigenomics. Differences in DNA methylation and histone acetylation are associated with resistance to chemotherapy; the expression of certain genes involved in the pharmacogenomics and pharmacodynamics of a therapeutic is modulated. [15] Such a change could not only impact the efficacy of a drug but also potentially lead to altered safety.

Understanding the disease biology at the patient level could therefore have major implications on treatment regimens where a drug and the dose are customized based on known factors contributing to a disease. Furthermore, a comprehensive understanding of biomarkers used for monitoring or predicting treatment toxicity could help guide clinicians through this process. Figure 6 shows a select list of biomarkers used to predict treatment toxicity in breast cancer.

Figure 6: Select list of biomarkers used to predict treatment toxicity in breast cancer.

Biomarker name	Biomarker type	Highest validity	Role	Technique	Substrate	Validity
Carbonyl reductase [NADPH] 1	Genomic	Late Studies in Humans	Predicting Treatment Toxicity	PCR + DirectSeq (DNA)	DNA	Early Studies in Humans
Dihydropyrimidine dehydrogenase [NADP+]	Proteomic; Genomic	Recommended / Approved	Predicting Treatment Toxicity	Genotyping (DNA)	DNA	Late Studies in Humans
Dimethylaniline monooxygenase [N-oxide-forming] 2	Proteomic; Genomic	Late Studies in Humans	Predicting Treatment Toxicity	Real Time PCR (DNA)	DNA	Early Studies in Humans
Histamine N-methyltransferase	Proteomic; Genomic	Late Studies in Humans	Predicting Treatment Toxicity	Real Time PCR (DNA)	DNA	Early Studies in Humans
Hyaluronan synthase 3	Proteomic; Genomic	Late Studies in Humans	Predicting Treatment Toxicity	PCR + DirectSeq (DNA)	DNA	Early Studies in Humans

Source: Cortellis Drug Discovery Intelligence

Role of artificial intelligence (AI)

A recurring theme in the previous sections is the advancement in technology that has provided the ability to quickly test multiple hypotheses and armed us with the necessary data to make informed decisions. Access to such large quantities of experimental data is crucial to gain a deeper understanding of the mechanisms related to toxicity and subtle differences in the genetic makeup that influence the maximum tolerated dose or efficacy of drug in different individuals. The next challenge therefore is quickly and efficiently managing and processing the mountains of data in a way that is meaningful and free of human bias. For this precise reason, there is an obvious need to incorporate AI and machine learning algorithms into data analysis workflows. To accomplish this, we need to understand the common types of data for which AI could be useful in analyzing and the potential approaches to handling them.

OMICs data

In the context of toxicology, OMICs data analysis would primarily be geared towards identifying biomarkers and constructing AoPs. This would be facilitated through computational methodologies that prioritize and validate the genes involved in adverse outcomes. A crucial component of this approach is consideration of the prior knowledge of the interactome (i.e, known interactions amongst the genes in an organism, including the genes being sampled and their neighborhood) and known associations of the members of the interactome and their mechanistic association to toxicity. Knowledge of protein-protein interaction networks is increasingly being used to determine drug-related side effects, and some of these examples have been previously reviewed and presented.[16]

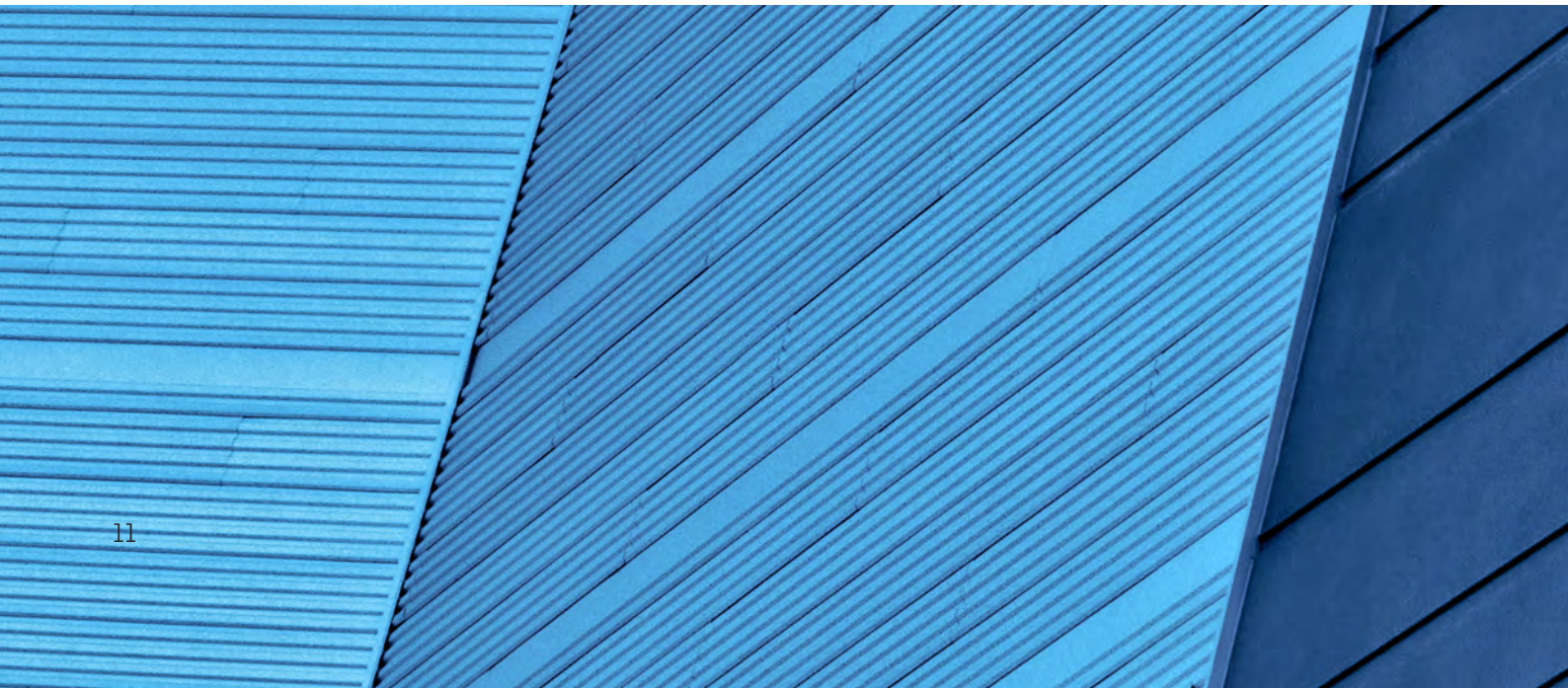
High-throughput screens

The datasets generated from high-throughput screening rely heavily on computational methods to crunch numbers to generate dose-response curves and predict the median lethal dose, or LC50, for a set of compounds. However, there is also growing demand for machine learning methods to tackle image recognition and classification in context to toxicology. Machine learning tools have been successfully used to understand the tissue pathology and to distinguish cancerous from non-cancerous regions.[17] Therefore, a similar concept can be used for image-based cardiotoxicity screening bioassays to predict dose-dependent cardiotoxicity.[18]

Pharmacokinetics and pharmacodynamics modeling

While ADME/Tox QSAR models have been around for a few decades, machine learning of pharmacokinetics data was used for the first time to predict patient-specific irinotecan toxicity in patients with metastatic colorectal cancer.[19] Anthropomorphic, biochemical and analytical data from individual patients were used to predict drug-induced side effects with 91% accuracy. Such predictions can have huge implications for personalized medicine, where fine tuning of pharmacokinetics factors based on patient data could significantly diminish adverse events.

Anthropomorphic, biochemical and analytical data from individual patients were used to predict drug-induced side effects with 91% accuracy.



AI challenges

While AI offers an exciting solution to our ever-growing problem of analyzing big data, we need to err on the side of caution. As with any new technology, there are some challenges associated with developing AI.

High-quality data

For machine learning, high-quality training datasets are needed to make accurate predictions. Whether it is ADMETox prediction models such as the ability of a compound to block the hERG channel or predicting tissue pathology based on image recognition, the quality of underlying data is very important to determining prediction accuracy. Similarly, other computational algorithms used for target deconvolution and biomarker prediction rely on the quality and granularity of prior published knowledge of the interactome, disease and toxicity associations. Thus, it is critical to source the information from a reliable and experimentally validated source.

Another major challenge for machine learning is the need for large quantities of datapoints to be fed into the model in order to achieve a high level of accuracy. Because of this, developers often turn to public data sources with the caveat that the data sourced may not be a true representation of the biological or experimental condition being modeled. A variety of public and commercial databases offer curated information on drug pipelines, and these could be important sources of information for machine learning. A recent study on HIV-1 integrase, protease and reverse transcriptase inhibitors illustrates the utility of these types of resources for developing QSAR models.[20]

Reproducibility

The use of computational techniques to solve biological problems is a relatively young field. To be incorporated into routine drug discovery workflow, these methodologies need to be reproduced and validated. Citing the lack of reproducibility and existing guidelines for reporting and publishing new algorithms, a recent paper from the Allen Institute for Artificial Intelligence developed a checklist for reporting experimental results; the checklist includes providing details on computational infrastructure, training sets and performance, implemented code and average run-time.[21] Implementation of this type of strategy will help streamline the use and validity of computational techniques in the future.

Looking forward to a ‘safe’ future

The ballooning cost of human healthcare has forced the question of whether the current practices in the healthcare industry are even sustainable. As the cost of bringing a drug to the market increases each year, we need to take a hard look at the underlying factors contributing to these numbers. Drug-induced toxicity and adverse events figure prominently at the core of these issues. In this review, we have looked at the best practices that are being implemented at various levels to streamline the process and make it more cost-effective.

A central theme that has emerged from this evaluation is the increased reliance on big data and the need for multiple testing to enhance the accuracy of adverse event predictions. As a result, there is a growing dependence on AI to quickly and reliably sift through the volumes of data and make accurate predictions of drug-induced toxicity and a patient’s susceptibility based on age, ethnicity, gender, and other characteristics. The future of economically viable drug development programs that deliver safe and efficacious therapies lies in changing our approach towards toxicology. No matter at which stage of the process our expertise lies, by implementing the strategies outlined in this paper, we can collectively be a part of the change.

References and background reading

1. businesswire.com/news/home/20171026006058/en/Top-3-Drivers-Toxicology-Testing-Market-Technavio. Accessed.

2. In. *CMR international pharmaceutical R&D factbook. (2019)*.

3. Matthews H, Hanison J, Nirmalan N. "Omics"-Informed Drug and Biomarker Discovery: Opportunities, Challenges and Future Perspectives. *Proteomes*. 2016;4(3).

4. Holmgren G, Sartipy P, Andersson C, Lindahl A, Synnergren J. Expression Profiling of Human Pluripotent Stem Cell-Derived Cardiomyocytes Exposed to Doxorubicin-Integration and Visualization of Multi-Omics Data. *Toxicological Sciences*. 2018;163(1):182-195.

5. Holmgren G, Synnergren J, Bogestål Y, et al. Identification of novel biomarkers for doxorubicin-induced toxicity in human cardiomyocytes derived from pluripotent stem cells. *Toxicology*. 2015;328:102-111.

6. Jiang J, Pieterman CD, Ertaylan G, Peeters RLM, de Kok TMCM. The application of omics-based human liver platforms for investigating the mechanism of drug-induced hepatotoxicity in vitro. *Arch Toxicol*. 2019;93(11):3067-3098.

7. Xie L, Evangelidis T, Bourne PE. Drug discovery using chemical systems biology: weak inhibition of multiple kinases may contribute to the anti-cancer effect of nelfinavir. *PLoS Comput Biol*. 2011;7(4):e1002037.

8. Xie L, Li J, Bourne PE. Drug discovery using chemical systems biology: identification of the protein-ligand binding network to explain the side effects of CETP inhibitors. *PLoS Comput Biol*. 2009;5(5):e1000387.

9. Ang ML, Murima P, Pethe K. Next-generation antimicrobials: from chemical biology to first-in-class drugs. *Arch Pharm Res*. 2015;38(9):1702-1717.

10. Cha Y, Erez T, Reynolds IJ, et al. Drug repurposing from the perspective of pharmaceutical companies. *Br J Pharmacol*. 2018;175(2):168-180.

11. Leist M, Hasiwa N, Rovida C, et al. Consensus report on the future of animal-free systemic toxicity testing. *ALTEX*. 2014;31(3):341-356.

12. Hartung T. Utility of the adverse outcome pathway concept in drug development. *Expert Opin Drug Metab Toxicol*. 2017;13(1):1-3.

13. Cochet A, Quilichini G, Dygai-Cochet I, et al. Baseline diastolic dysfunction as a predictive factor of trastuzumab-mediated cardiotoxicity after adjuvant anthracycline therapy in breast cancer. *Breast Cancer Res Treat*. 2011;130(3):845-854.

14. Schneider BP, Shen F, Gardner L, et al. Genome-Wide Association Study for Anthracycline-Induced Congestive Heart Failure. *Clin Cancer Res*. 2017;23(1):43-51.

15. Monte AA, Vasiliou V, Heard KJ. Omics Screening for Pharmaceutical Efficacy and Safety in Clinical Practice. *J Pharmacogenomics Pharmacoproteomics*. 2012;S5.

16. Kotlyar M, Fortney K, Jurisica I. Network-based characterization of drug-regulated genes, drug targets, and toxicity. *Methods*. 2012;57(4):499-507.

17. Dabeer S, Khan M, Islam S. Cancer diagnosis in histopathological image: CNN based approach. In. Vol Volume 16, 2019, 100231. Informatics in Medicine Unlocked: Elsevier; 2019.

18. Joshi-Barr S, Haberland A, Bartel S, Müller J, Choi T, Wallukat G. High throughput bioassay for beta1-adrenoceptor autoantibody detection. *Int J Cardiol*. 2016;219:98-104.

19. Oyaga-Iriarte E, Insausti A, Sayar O, Aldaz A. Prediction of irinotecan toxicity in metastatic colorectal cancer patients based on machine learning models with pharmacokinetic parameters. *J Pharmacol Sci*. 2019;140(1):20-25.

20. Stolbov LA, Druzhilovskiy DS, Filimonov DA, Nicklaus MC, Poroikov VV. (Q)SAR Models of HIV-1 Protein Inhibition by Drug-Like Compounds. *Molecules*. 2019;25(1).

21. Dodge J, Gururangan S, Card D, Schwartz R, Smith NA. Show Your Work: Improved Reporting of Experimental Results. In. arXiv:1909.03004 [cs.LG]2019.

About Cortellis

Cortellis™, a Clarivate Analytics solution, gives life to science by unlocking the hidden insights in data by curating broad and deep sources of intelligence to enable precise, actionable answers to specific questions across the R&D lifecycle — from discovery and clinical development through regulatory submission and commercialization. By supporting data-driven decisions, Cortellis helps pharmaceutical companies, biotech and medical device/diagnostic firms accelerate innovation. Over the past year, 80% of U.S. companies filing NMEs, 91% of companies achieving breakthrough therapy status and 70% of the top licensing deals were informed by Cortellis intelligence. For more information, please visit clarivate.com/cortellis.

clarivate.com/cortellis

LS479028910 / 01