Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about *Accepted Manuscripts* in the **Information for Authors**.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard <u>Terms & Conditions</u> and the <u>Ethical guidelines</u> still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.



www.rsc.org/toxicology

Translation of off-target effects: Prediction of ADRs by integrated experimental and computational approach

Laszlo Urban^{1*}, Mateusz Maciejewski¹, Eugen Lounkine¹, Steven Whitebread¹, Jeremy L. Jenkins², Jacques Hamon³, Alexander Fekete¹ and Patrick Y Muller⁴

¹Preclinical safety Profiling, Center for Proteomic Chemistry, Novartis Institutes for Biomedical Research, Cambridge, Ma., USA; Developmental and Molecular Pathways, Novartis Institutes for Biomedical Research, Cambridge, Ma., USA; ³Basel Screening Operations, Center for Proteomic Chemistry, Novartis Institutes for Biomedical Research, Basel, Switzerland, Center for Proteomic Chemistry, ⁴Global Pharma Development Strategy, Novartis Pharma, Basel, Switzerland

Toxicology Research Accepted Manuscript

Abstract

Adverse drug reactions (ADRs) are associated with most drugs, often discovered late in drug development and sometimes only during extended course of clinical use. They are linked either to the therapeutic target or pathway, or could emerge as the consequence of known or unknown off-target effect(s) of a drug or drug combinations. ADRs are a major burden on patients, medical professionals and the society in general. Discovery of intolerable ADRs during clinical trials significantly contributes to high attrition rates with associated rising cost. Thus, prediction of ADRs at the early stage of drug discovery is an emerging approach; however, it remains a challenging task to identify the mode of action of drug candidates which might lead to ADRs. We review here the implementation of *in vitro* and *in silico* tools streamlined for the prediction of ADRs as early as the target/lead identification and lead optimization phases of the drug discovery process. This integrated approach has been developed during the past decade by both academic institutions and the pharmaceutical industry with the aim to provide toxicological analysis, assessment and ranking of drug candidates on a broad scale. The major aim is to be able to mitigate targets associated with ADRs earlier and guide chemistry to address the therapeutic and side effects in parallel. The maior components of this effort are (1) experimental approach: early in vitro safety profiling linked to (2) computational toxicology algorithms and models utilizing statistics, data mining, cheminformatics and system biology. The third component embraces the translational aspect for clinical ADRs, which includes in vivo exposure. In this review we focus on the prediction of the integrated molecular network approach.

Page 3 of 35

The burden of adverse drug reactions

"A main factor contributing to increased drug-related expenses has roots in early drug discovery; the cost, however, will appear in the clinic"¹. Namely, the majority of safety issues manifest during clinical trials and post-approval. Historically, safety assessment of drug candidates was left to the end of the drug discovery process, when it contributed Go/NoGo decisions, almost exclusively based on *in vivo* safety pharmacological observations and histopathological data. Hazard identification in parallel with optimization for potency and surrogate biological effect was almost completely non-existent. This "purist" approach has resulted in large-scale attrition rate during clinical trials and with drugs entering the clinic without knowledge of therapeutic target or off-target related side effects, leading to several withdrawals and numerous box labels². It is not a surprise that side effects of individual drugs or drug combinations generate a significant expense in healthcare and early safety assessment has been introduced by the pharmaceutical industry and more stringent rules are demanded by the regulatory authorities.

Side effects, or adverse drug reactions (ADRs) can result from either non-therapeutic effects associated with the primary target, e.g. in a different organ (e.g., CNS effects of H1 antihistamines³), or by unintended effects on "off-targets" such as inhibition of the human ether a go-go related gene (hERG) potassium channel resulting in proarrhythmic QT prolongation⁴. It is important to note, that any protein could become off-target when engaged unintentionally and can generate side effects. As an example, inhibition of the 5-HT3 channel could be a therapeutic target for narcolepsy, however blocked as an off-target could cause prolonged wakefulness and insomnia.

Toxicology Research Accepted Manuscript

Side effects could develop acutely, concomitant with the therapeutic effect of a drug or as a delayed chronic ADR, such as was observed following treatment of breast cancer with tyrosine kinase inhibitors (TKIs)^{5,6}. ADRs are however complex phenomena, which can have multiple factors, such as pathway perturbations, pharmacokinetic components, drug-drug-interaction issues, metabolic components and formation of reactive metabolites. In this review, we will focus on the possible prediction of on- and off-target effects, using a practical combination of early *in vitro* assessment and computational models and algorithms that aim at establishing a link between molecules, their biological (off)-targets, and ADRs.

Up to date cardiac and hepatic toxicities have been the leading reasons for attrition and labelling⁷. These side effects largely develop because of off-target activity of drugs and their metabolites. Even though prediction of pro-arrhythmic potential of drugs has improved considerably, other aspects of cardiotoxicity, such as development of cardiomyopathies or valvular heart disease, significantly contributed to cardiovascular side effects^{8,9}. Idiosyncratic drug-induced liver injury (DILI) is another leading ADR with various underlying mechanisms. Recently, newly developed *in vitro* and *in silico* technologies have been implemented to assess drug-drug interactions, toxic effects of reactive metabolites and link hepatobiliary hepatotoxicity to bile transporters^{10,11}.

The emerging landscape is that as we learn more about mechanisms of action underlying toxic effects, the term "idiosyncratic" is removed from an increasing number of ADRs. This is true not just for hepatic and cardiovascular toxicity, but in general for all organ-linked ADRs¹². Importantly, we have learnt to address mitigation of off-target effects and producing drugs with less pharmacological promiscuity^{13,14}, in particular

concerning "antitargets"^{12,14,15}. For example, preclinical assessment of cardiac ion channel inhibition can support a well-defined mitigation process prior to candidate selection and prevents compounds, with potent effects at these channels, entering clinical trials. Consensus within the pharmaceutical industry concerning target-based profiling practices was covered by recent reviews^{12,16}.

Early awareness of safety hazards

Increasing regulatory requirements contributed to extensive preclinical safety assessment of compounds at an earlier stage of drug discovery¹⁷. Clinical trials involving broader patient populations in multiple centers and inclusion of expensive and time consuming safety related procedures are more common requirements^{18,19,20}. It is in interest of the pharmaceutical industry to be prepared for such investigations, thus early assessment of hazards associated with molecular targets and pathways are agining ground and enter the domain of registration expectations^{18,19}. As an added benefit, data accumulated by in vitro safety profiling opened the gate to the development of *in silico* tools with predictive power for ADRs^{21,22}. This stems from two aspects of early safety assessment: target molecule-ADR associations became known. however the experimental approach of testing a large set of molecules in a very broad pharmacology space is neither technically nor financially feasible. Thus, linking "chemical space" to ADRs via molecular networks using comparably inexpensive in silico models, which can be applied on a large scale, represents a viable approach to fill this gap.

Introduction of secondary pharmacology to support early safety assessment

Adverse drug reactions rarely appear as a single symptom or manifest in one particular biomarker. In practice, most drugs are associated with a selection of adverse events, creating a characteristic phenotype. This depends on the therapeutic and off-target profile of the drug, access to the different organs and patient population. However, to find ways to mitigate ADRs during the drug discovery process, we need to identify individual targets associated with particular side effects. This is where in vitro secondary pharmacology comes into scope^{12,17}. It can be done at the beginning of the drug discovery process, in parallel with lead selection and optimization at a low cost. However, this approach only works if data generated in these assays or assay panels are predictive for well-defined ADRs in humans. Therefore, reverse translation is used to link ADRs to targets/pathways for the identification of targets to be included into the target-based safety profiling panels. Some on-target ADRs of drugs are relatively easy to identify based on their intended, known pharmacology. For example, PDE3 inhibition is well known to increase the incidence of death in congestive heart failure patients²³ because of longer-term consequences of its intended positive inotropic effect. The diagram using reverse translation for identifying target-related ADRs is presented in Figure 1. This approach is based on the assumption that identical ADRs of medicines for diverse indications are likely to be associated with the same target or a close target in the same pathway. The schematic diagram highlights in orange the crucial step where the connection between the ADR and target is established. While this has been achieved in the past to great extent by experimental approach based on empirical data, development of novel in silico methods applying Bayesian models^{24,25} and molecular

network approach^{21,26,27} offer predictions of novel drug-target-ADR associations as will be discussed later.

In the absence of drugs in the clinic, OMIM, animal data from KO/KI experiments and *in vivo* pharmacology will provide information on targets and associated phenotypes which help to recognize ADRs in the clinical settings. Target-ADR associations based on animal data have lesser confidence value because translation for clinical ADRs will depend on species specificity of targets, relative potency of compounds, pharmacokinetic factors and differences in animal and human phenotypes associated with the same or closely homologous proteins.

Let's consider the benefits and limitations of *in vitro* secondary pharmacology. It is important to emphasize that it provides alerts for safety hazards by identifying off-targets. *In vitro* secondary pharmacology assessment could fall short in the prediction of clinical manifestation of the off-target effects. It is relatively easy to predict unwanted off-target effects; however, their manifestation and translation to ADRs is determined by the complex pharmacodynamics/pharmacokinetic performance of the drug candidate under clinical conditions. The potency at the off-targets should be considered in the context of exposure, which is significantly defined by the maximum free available drug concentration at the site of therapeutic action in the organism²⁸. Most often the free C_{max} is used for this purpose, but AUC can be considered if more relevant. In general, this simple calculation will forecast whether the hazard would represent a safety risk during clinical application. Once the risk has been established, *in vitro* profiling assays can lead mitigation during lead optimization by defining and applying structure-activity relationship (SAR), and guide the preclinical safety assessment in finding the "no

observed adverse effect level" (NOAEL) for the associated ADR in case the hazard still persists at candidate selection. To demonstrate the practical use of this concept, we consider the case of 5-HT2B (HTR2B) agonism associated with the development of cardiac valvular disease⁸: If a clinical candidate with high potency at the therapeutic target shows weak agonist activity at this off-target then a long-term rodent experiment including both echocardiography and histopathology should be considered within the preclinical assessment²⁹ for the support of a final decision for further development of the compound.

A valuable application of *in vitro* safety pharmacology assessment involves preclinical prediction of hazard and the associated risk for such particularly difficult, multifactorial ADRs as suicidal ideation. The exclusively human aspect of this ADR lacks relevant *in vivo* evaluation. However, once critical evidence emerges from clinical observations - in particular from the FDA adverse event recording system (FAERS³⁰) - on high incidence of suicidal ideation of various drugs with common central nervous system (CNS) targets and/or off-targets, one can establish links between the target and the observed ADR. Testing compounds at these targets will give an opportunity for early warning for suicidal ideation, support mitigation and initiate *in vivo* testing guided by the knowledge of the suspect off-target.

Pharmacological promiscuity and poly-pharmacology

Pharmacological promiscuity is considered an important component in side effect prediction^{13,14}. There are two aspects to be looked at: promiscuous ligands¹³ and

promiscuous targets³¹. While well composed target panels will consider both of these components, they only cover a fairly small segment of pharmacological space, whether they are solely proprietary, or include existing data published in literature and deposited in chemogenomic databases. Thus the full scale of promiscuity will remain hidden with the consequence of possible unexpected ADRs. For practical purpose, relatively slim profiling panels of targets that received high scores for seriousness of associated ADRs and also for high hit rate by compounds in the test set (e.g., promiscuous targets, such as kinases) are implemented for regular, iterative safety assessment^{12,17}. However, testing a large number of compounds for safety assessment in a very broad *in vitro* panel would be expensive and time consuming. To address this bottleneck, *in silico* models have become increasingly popular as they can cover broad chemical and biological space.

In certain cases poly-pharmacology is a desirable feature. One might want to alter the activity of multiple targets that together result in the desired phenotype. This task is hard to achieve and integrate with safety profiling using traditional medicinal chemistry considerations, and in practice can be reliably accomplished only with the use of computational³² modeling. Models vary between approaches with focus on chemical or biological characteristics with best results when both are considered and integrated.

Modeling safety aspects of drugs: Integrated molecular network approach

Alignment of ontologies provide a solid framework for computational toxicology

During the past 20 years there has been a significant increase of data collection with the advent of automatic screening and IT technologies. Large preclinical datasets, generated on a vast number of molecules in a diverse set of in vitro assays and collection of clinical data associated with drugs opened up opportunities for model building. While these advances have provided foundation for *in silico* approaches, the lack of consistent annotation, coherent ontology, and non-standard terminology created obstacles against fast advance. Non-uniform ontologies prove to be one of the main factors complicating the task of data mining across multiple databases. Despite extensive standardization efforts³³, discrepant coding, in particular of drug names and ingredients, remain a major challenge in preclinical assessment. Safety related assessment often follows Medicinal Dictionary for Regulatory Activities (MedDRA³⁴) terms and relate to the FDA Adverse Event Reporting System (FAERS³⁰) which provides the translational bridge between preclinical and clinical information. Both Elsevier's Pharmapendium³⁵ and Thomson Reuters Integrity³⁶ are utilizing these resources and provide useful sources of information on marketed drugs. Despite its availability and easy public access, FAERS should be approached with caution, as multiple normalization steps are required to be done by users. Specifically, although ADRs in FAERS are expressed using the MedDRA ontology, the names of drugs in the reports are entered in free form and are not standardized. This complicates the monitoring of adverse events related to individual ingredients and can lead to false negatives in the analysis of drug – adverse drug reaction signals in this database.

Additionally, target-ADR associations can be deduced from known gene-disease associations. The Online Mendelian Inheritance in Man (OMIM) database³⁷ is an

10

Toxicology Research Accepted Manuscript

example resource that can be used to predict novel ADR effects associated with interaction with specific targets. While OMIM was designed for human (rather than computer) readers, and as such does not support any ontology, some attempts have been undertaken to translate the OMIM terms into MeSH ontology³⁷, and such translated versions of OMIM are available in some databases, such as Thomson Reuters Integrity³⁶.

Also, a large volume of data concerning safety assessments of compounds is hidden from public access in company databases. To rectify this problem regulatory agencies, academia and industry joined resources (see coordinated efforts from EPA, NIH, FDA and EMEA) by generating projects on a large volume of compounds and make them available for open access^{38,39,40}.

In silico models to predict pharmacological profiles (phenotypes) of compounds

Small molecule drug discovery is largely based on the application of structure activity relationship (SAR), with the assumption that similar molecular structures have similar biological activities, often referred to as the similarity property principle. This concept is primarily used for lead optimization to improve activity at the indication-specific therapeutic target. However, this approach can be applied to the same extent to off-targets, since a molecule doesn't "care" about its intended action. If the high-resolution target structure is known, docking can be a valuable *in silico* profiling approach^{41,42}.

Prior to addressing molecular networks associated with toxicological processes, phenotypes, we will discuss molecular similarity and statistical cheminformatics approaches which support more complex in silico methods and can be applied early in

Toxicology Research Accepted Manuscript

the lead discovery pipeline to prioritize sets of compounds. These models opened up the way to interpret ADRs by connecting drugs via a network of targets to ADRs. They rely on the similarity property principle and utilize the increasing pharmacological knowledge gained through profiling efforts, including proprietary and public sources. The basic underlying idea is that compounds similar in their structure to known ligands of a target are likely to modulate that target as well. These models provide qualitative predictions (likely to bind / not likely to bind). Despite this limitation, they can be readily applied to large and diverse compound sets, which can support early safety assessment for hazard identification. Thus, they are suited to prioritize compounds for screening and scaffold selection for further optimization, rather than to make decisions on individual compounds.

Although the underlying principle is the same for many of these models, they can be organized based on the level of abstraction that they employ. Assume we want to predict off-targets of a test compound. Nearest-neighbor (NN) approaches extend the idea of an analogue search by assessing the overlap of unique chemical features, and thus quantifying molecular similarity. Thus, if the test compound is highly similar to a known off-target ligand, it is predicted to modulate that target as well (Figure 2). Building on this idea, (off-)targets are represented by a collection of their known ligands. The similarity of the test compound can be expressed as a simple arithmetic mean (n-NN), or an average of the most similar compounds (k-NN).

Evolving this idea, Keiser et al.^{43,44,45} have developed the similarity ensemble approach (SEA), which assigns statistical significance to this combined similarity by comparing it to the combined similarity of a random test compound to a random set of ligands the

size of the target set. This approach has been applied to predict off-targets of marketed drugs on a large scale ²¹. Because it takes into account the similarity to the entire set of known ligands, each individual, pair-wise similarity can be rather small, but the sum can nevertheless be significantly higher than expected by chance. This advantage is reflected in the often surprising prediction for compounds that show only little structural resemblance to any single known ligand of a target. Based on the above, these approach was found suitable for testing compounds in silico for unexpected (e.g. off-) targets during lead optimization. While determination of affinity remains out of scope of this method, it is a reliable tool for hazard identification within a scaffold or for single molecules in conjunction with confirmation of the prediction in secondary pharmacology assays²¹.

The next step of abstraction is statistical models, such as naive Bayes, in which chemical features overrepresented among known ligands are used individually to score the test compound. This approach is highly resistant to noise in the training data, and profits from a large collection of diverse target ligands for training⁴⁶.

In order to link the predicted targets to ADRs, additional data can be utilized, such as gene expression data⁴⁷ and pharmacokinetic information which allows to identify biological pathways and targets associated with complex ADRs^{24,48}. In particular, drug-target-ADR networks employ a guilt-by-association metric to assess whether an observed ADR is more likely due to the primary or another well-known target, and which adverse events may be attributed readily to novel predicted off-targets²¹.

So far we discussed how molecular similarity and statistical approaches can be applied to describe targets using their ligands. However, considering sets of drugs with known

ADRs, these adverse events can also be represented by chemical features. Bender et al.²⁵ established a proof-of-concept for predicting adverse events based on chemical structures alone by linking targets to ADRs through correlation in chemical feature space. For example, chemical features statistically overrepresented for bioactivity against a target in large-scale bioactivity databases may also be overrepresented among drugs that cause the same side effect (Figure 3). Statistical models (e.g. Bayesian models) can be directly compared by computing the correlation of weights placed on chemical features between two models (either target-target, ADR-ADR, or target-ADR models). These correlations are usefully visualized in correlation networks which can reveal groups of targets and ADRs that cluster based on chemical features associated with them. This line of thinking has been further developed by combining ADR and target associations with chemical structures together with gene-pathway annotations^{24,25}. The main hypothesis is that compounds which hit different targets in the same pathway can cause the same phenotypic effect. By aggregating compounds that share the same ADR phenotype and applying *in silico* target prediction, pathways causing ADRs can be revealed. We have referred to this aspect of the above in silico methods in the introduction and highlighted their use for target or MoA identification in a reverse translation mode. In our practice, this approach plays an important role to select secondary pharmacology targets as well.

Efforts on *in silico* ADR prediction are not isolated. There are several initiatives, including the European Innovative Medicines Initiative (IMI) eTOX (expert systems in Toxicology) Consortium⁴⁹ which mines preclinical and clinical trial information never published by pharmaceutical companies to reveal hidden associations between

chemical structures, target molecules/pathways and ADRs. Several US federal agencies (National Toxicology Program/ National Institute of Environmental Health Science, National Center for Advancing Translational Sciences and the Food and Drug Administration) also joined forces to undertake a major screening project called Toxicity Testing in the 21st Century (Tox21)³⁸ to explore the *in vitro* effects of drugs, environmental toxins and industrial/household chemicals on a large number of targets, scattered in the pharmacology space. Within the frame of this massive collaboration about 10,000 compounds are being tested. ToxCast results generated by EPA^{39,40}are contributed to the federal agency collaboration and are open for the public⁴⁰ through user-friendly web applications called "interactive Chemical Safety for Sustainability Dashboards (iCSS). These projects aim to provide open access to extensive biological and pharmacological profiling of compounds the public is exposed to through mining tools linking the data to observed effects.

Similarly, other applications of parallel pharmacophore screening have been published recently^{50,51}. In fact, Oprea et al. called for an in-depth integration of these sources in newly developed cheminformatics tools and coined the fitting term "Systems Chemical Biology"⁵². In a first step, Scheiber et. al²⁵ employed a large set of ligand-based protein target prediction models (>2000) generated using chemical fingerprints for compounds from GVK Bio databases,⁵³ WOMBAT,⁵⁴ MDDR,⁵⁵ and in-house Novartis databases to automate ligand target prediction. Importantly, all targets were mapped to the same ontology (e.g., NCBI gene symbol) to enable aggregation of bioactivity data from different sources. Targets which were predicted for a set of compounds sharing the same ADR were mapped into pathways in MetaBase⁵⁶ (GeneGo/Thomson Reuters) to

Toxicology Research Accepted Manuscript

i age i o o i

retrieve those that contain at least one of the predicted targets. Finally, pathways were ranked according to their most relevant link to the selected ADR.

As computational models help to expand predictions for biological and toxicological effects of small molecules there is a need to establish direct link with the underlying pathways and mode of action. A good example is the development of BioMAP⁵⁷ and several other models amalgamating in silico and "organ-on-the-chip" technologies^{58,59}. System biology using human cells or tissues is often applied for target prediction however this approach utilizes both in a very close conjunction. Various chips model human conditions based on microfluidics and compartmentalized organ specific tissues, cell cultures, biopsies or more often stem cells or iPS cells derived from patients with genetic diseases. This approach is most promising to provide early human data, including metabolism and pathophysiological environment resembling disease phenotypes^{58,59}. The BioMAP model was initiated to overcome the complexity of proteomic microarrays which are difficult to interpret. It is constructed of groups of primary human cell co-cultures with close resemblance of their physiological assembly, stimulated with biologically relevant combinations of inflammatory mediators which will provide a specific phenotype based on a limited number of highly characteristic redouts. Modulation of selected protein readouts by chemical agents are presented as activity profiles (BioMAP profiles). Statistical methods are applied to the profiles to identify similarities between patterns linked to mode of action prediction. The close amalgamation of in vitro system biology with in silico tools provides a remarkable system capable of identifying associations of pathways with drug biological read-outs. For details on BioMAP consult publications by Plavec et al.⁶⁰ and Storey & Tibshirani⁶¹.

A recent application of BioMAP within the ToxCast program successfully proved that drug-pathway pairing can be associated with toxicological liabilities and related ADRs⁶². System biology approaches have the power of providing direct human data with tissue or even organ-specific format, with strong "disease phenotypic" aspects and could identify biomarkers for further use in clinical trials.

Many *in silico* approaches towards therapeutic and safety assessment use network considerations^{26,63,64,65}, therefore it is no surprise that they have converged in recent vears into a scientific discipline of its own. The term "network medicine"²⁶ was coined to describe this field of research which includes the various methodological approaches, with an emphasis on the "interactome" and the general organizing principles that govern cellular networks. The human interactome is a network that includes all known molecular interactions in the human organism. It is divided (anthropomorphically) into sub-networks: protein-protein interactions, drug/ligand/substrate-protein interactions, protein nucleic acid interactions (regulatory, RNA, and metabolic coupling). Modules or sub-networks could be identified within such interactomes that reasonably represent disease states^{26,63}. In the context of drug safety, drug-target networks emphasize drugmultiple target–pathway (or systems biology) based phenotypes⁶³. Linking interactome data with cheminformatics and computational chemistry models provides a promising approach to computational toxicology by better understanding drug function through disease gene associations and interconnectedness between cellular pathways. Networks can model perturbed disease states by symptomatic treatment and identify accurate biomarkers²⁶. Historically, drug side effects were not easy to access for data mining purposes. The introduction of the side-effect database, SIDER⁶⁶

(http://sideeffects.embl.de) helped to advance side-effect network studies²⁷. Lee et al.⁶⁷ extended the SIDER-derived side-effect networks by biological processes (Gene Ontology terms and PubMed data text mining). Interestingly, when SIDER data were combined with information about disease-associated genes, it was observed that drugs hitting targets 3 or more edges away in a network from disease genes yielded fewer side effects⁶⁸. This method can be considered for rational drug design to balance therapeutic – ADR effects.

Utilization of the human interactome offers a new perspective for safety assessment, namely the in silico application of chemical structures to perturbed pathways which is very rarely achieved in preclinical safety assessment models. One can expect significant refinement of the human interactome with accumulating information on molecular and developmental pathways and ADR phenotypes with the benefit of improving efficacy of in silico tools which are based on this approach.

Conclusions

The development of novel *in vitro* technologies in association with the emerging *in silico* methods have provided a new type of safety assessment for small molecule drug candidates, pushing the frontline further upstream in the drug discovery process. The classic, largely regulatory toxicological evaluation of single compounds has been complemented by an early, integrated chemical and biological network-based safety assessment. These models are aided by links to clinical data, using translational information obtained from ADRs. The connection of all of these components could not be achieved without computational methods, including large scale search capabilities and smart algorithms enabling cheminformatics and system biology evaluation of ever

increasing volume of biological and chemical data. Ligand- and pathway-based in vitro safety profiling efforts are now closely integrated with *in silico* methods and provide decision support tools for safety assessment in conjunction with efficacy information. Databases generated and enhanced by academic, industrial and government institutions obtained from target- and pathway-based safety profiling assays provide a rich collection of reference data more often with synchronized annotation and recently with the same ontology. This enables cross searches of large, high quality databases for machine learning tools and statistical models. Some of the models and databases are readily accessible for the scientific public, others are commercially available, or published as validated concepts.

In this review we focused on the prediction of integrated molecular network approach, while leaving discussion on models developed to address pharmacokinetics, organ specific toxicity and technical aspects to other previously published excellent reviews^{69,70,71}.

Safety profiling assays supported by the integrated molecular network approach are now well established and part of the drug discovery process. The extensive target annotation, association with ADRs and consideration of pharmacokinetic aspects provide confidence in the translational value of this method, thus it is more and more used for early safety risk assessment. Acknowledging, that the vast expanse of chemical and biological space does not allow "limitless" testing of compounds in the laboratory, *in silico* methods are more frequently used to rectify this shortcoming. Importantly, they also have a great utility in the design of new molecules devoid of off-

target effects and guide chemists to synthesize those structures which carry less or no hazard towards unwanted ADRs.

References

- L. Urban, J-P. Valentin, K. I. Kaitin, J. Wang, Current social, clinical and scientific environment of pharmaceutical R&D, In: Predictive ADMET: Integrated approaches in Drug Discovery and Development, Eds: J. Wang and L. Urban, *Wiley Inc.,* ISBN 978-1-118-29992-0, 2014, 3-22.
- K. E Lasser , P. D. Allen, S. J. Woolhandler, D. U. Himmelstein, S.M. Wolfe, D. H. Bor, Timing of new black box warnings and withdrawals for prescription medications, *JAMA*, 2002, **287**, 2215-2220.
- 3. J. P. Rihoux and F. Donnelly, CNS effects of histamine H1 antagonists, *Clin. Exp. Allergy.*, 1999, Suppl 3, 143-146.
- G. Gintant, An evaluation of hERG current assay performance: Translating preclinical safety studies to clinical QT prolongation, *Pharmacol. & Therapeutics*, 2011, **129**, 109–119.
- K. H. Schmitz, R. G. Prosnitz, A. L. Schwartz, J. R. Carver, Prospective surveillance and management of cardiac toxicity and health in breast cancer survivors, *Cancer*, 2012, **118**, 2270-2276.
- H. R. Mellor, A. R. Bell, J. P. Valentin and R. R. Robert, Cardiotoxicity associated with targeting kinase pathways in cancer, *Tox. Sci.*, 2011 **120**, 14-32.
- W. S. Redfern, J-P. Valentin, Trends in safety pharmacology: posters presented at the annual meetings of the Safety Pharmacology Society 2001-2010, *J. Pharmacol. Toxicol. Meth.*, 201, **64**, 102-110.
- 8. B. L. Roth, Drugs and Valvular Heart Disease, N. Engl. J. Med., 2007, 356, 6-9.

- T. Force and K. L. Kolaja, Cardiotoxicity of kinase inhibitors: the prediction and translation of preclinical models to clinical outcomes, *Nat. Rev. Drug Discov.*, 2011, **10**, 111-126.
- 10.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.
- 11.J. Wang Comprehensive Assessment of ADMET Risks in Drug Discovery, *Curr. Pharmaceut. Design*, 2009, **15**, 2195-2219.
- 12. J. Bowes, A. J. Brown, J. Hamon, W. Jarolimek, A. Sridhar, G. Waldron and S. Whitebread, Reducing safety-related drug attrition: the use of *in vitro* pharmacological profiling, *Nat. Rev. Drug Discov.*, 2012, **11**, 909-922.
- 13.K. Azzaoui, B. Faller, S. Whitebread, E. Jacoby, A. Bender, J. Jenkins and L. Urban, Modeling Promiscuity based on *in vitro* Safety Pharmacology Profiling Data, *Chem. Med. Chem.*, 2007, **2**, 874-880.
- 14. L. Urban, S. Whitebread, J. Hamon, D. Mikhailov and K. Azzaoui, Screening for Safety-relevant Off-target Activities, In Polypharmacology in Drug Discovery, 2012, Ed.: J-U. Peters, John Wiley and Sons, Inc.
- 15. R. J. Vaz and T. Clabunde (eds) Antitargets, 2008, Wiley-VCH.
- 16.L. Ewart, D.J. Gallacher, G. Gintant, J.M. Guillon, D. Leishman, P. Levesque, N. McMahon, L. Mylecraine, M. Sanders, W. Suter, R. Wallis and J-P. Valentin, How do the top 12 pharmaceutical companies operate safety pharmacology, *J. Pharmacol. Toxicol. Methods*, 2012, **66**, 66-70.

- 17. S. Whitebread, J. Hamon, D. Bojanic and L. Urban, *In vitro* Safety Pharmacology profiling: An essential Tool for Drug Development, *Drug Discov. Today*, 2005, 10, 1421-1433.
- 18. Anon. Guidance for Industry Diabetes Mellitus Evaluating Cardiovascular Risk in New Antidiabetic Therapies to Treat Type 2 Diabetes. U.S. Department of Health and Human Services. Food and Drug Administration Center for Drug Evaluation and Research (CDER), 2008,

http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation /Guidances/ucm071627.pdf

 Anon. Guidance for Industry Suicidal Ideation and Behavior: Prospective Assessment of Occurrence in Clinical Trials – Draft guidance. U.S. Department of Health and Human Services. Food and Drug Administration Center for Drug Evaluation and Research (CDER), 2012,

http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation

/Guidances/UCM225130.pdf

- 20. J. A. Kramer, Designing safe drugs: what to consider, *Expert Opin Drug Discov.*, 2008, **3**, 707-713.
- 21. E. Lounkine, M. J. Keiser, S. Whitebread, D. Mikhailov, J. Hamon, J. L. Jenkins, P. Lavan, E. Weber, A. K. Doak, S. Côté, B. K. Shoichet and L. Urban, Large-scale prediction and testing of drug activity on side-effect targets, *Nature* 2012, **486**, 361-367.
- 22. F. Nigsch, E. Lounkine, P. McCarren, B. Cornett, M. Glick, K. Azzaoui, L. Urban, P. Marc, A. Müller, F. Hahne, D. J. Heard and J. L. Jenkins, Computational biology

and chemistry incorporated into early predictive safety. *Curr. Opin. Comp.* Biol. 2011, **7**, 1497-1511.

- S. A. Aguirre, J. R. Heyen, W. Collette, W. Bobrowski and E. R. Blasi, Cardiovascular effects in rats following exposure to a receptor tyrosine kinase inhibitor, *Toxicol. Pathol.*, 2010, **38**, 416-428.
- J. Scheiber, B. Chen, M. Milik S. C. Sukuru, A. Bender, D. Mikhailov, S. Whitebread, J. Hamon, K. Azzaoui, L. Urban, M. Glick, J. W. Davies and J. L. Jenkins, Gaining insight into off-target mediated effects of drug candidates with a comprehensive systems chemical biology analysis, *J. Chem. Inf. Model.*, 2009, **49**, 308-317.
- A. Bender, J. Scheiber, M. Glick, J. W. Davies, K. Azzaoui, J. Hamon, L. Urban, S. Whitebread and J. L. Jenkins, Analysis of pharmacology data and the prediction of adverse drug reactions and off-target effects from chemical structure, *ChemMedChem.*, 2007, 2, 861–873.
- 26. A.L. Barabási, N. Gulbahce and J. Loscalzo, Network medicine: a network-based approach to human disease, *Nat Rev Genet.*, 2011, **12**, 56-68.
- P. Csermely, T. Korcsmáros, H.J. Kiss, G. London and R. Nussinov, Structure and dynamics of molecular networks: a novel paradigm of drug discovery: a comprehensive review, *Pharmacol. Ther.*, 2013, **138**, 333-408.
- P. Y. Muller and M. N. Milton, The determination and interpretation of the therapeutic index in drug development, *Nat. Rev. Drug Discov.*, 2012, **11**, 751-761.

- 29. S. Droogmans, P. R. Franken, C. Garbar, C. Weytjens, B. Cosyns, T. Lahoutte V. Caveliers, M. Pipeleers-Marichal, A. Bossuyt, D. Schoors, and G. Van Camp, *In vivo* model of drug-induced valvular heart disease in rats: pergolide-induced valvular heart disease demonstrated with echocardiography and correlation with pathology, *Eur. Heart J.*, 2007, **28**, 2156-2162.
- 30. Anon. FDA, http://www.fdable.com/basic_query/aers (accessed: 03. 15. 2014)
- V. I. Pérez-Nueno, V. Venkatraman, L. Mavridis and D. W. Ritchie, Detecting drug promiscuity using Gaussian ensemble screening, *J. Chem. Inf. Model.*, 2012, **52**, 1948-1961.
- J. Besnard, G. F. Ruda, V. Setola, K. Abecassis, R. M. Rodriguiz, X. P. Huang, S. Norval, M. F. Sassano, A. I. Shin, L. A. Webster, F. R. Simeons, L. Stojanovski, A. Prat, N. G. Seidah, D. B. Constam, G. R. Bickerton, K. D. Read, W. C. Wetsel, I. H. Gilbert, B. L. Roth and A. L. Hopkins, Automated design of . ligands to polypharmacological profiles. Nature, 2012, **492**, 215-220.
- 33. A. de Souza, J. A. Bittker, D. A. Lahr, S. Brudz, S. Chatwin, T. I. Oprea, A. Waller, J. J. Yang, N. Southall, R. Guha, S. C. Schurer, U. D. Vempati, M. R. Southern, E. S. Dawson, P.A. Clemonsand T. D. Chung, An Overview of the Challenges in Designing, Integrating, and Delivering BARD: A Public Chemical-Biology Resource and Query Portal for Multiple Organizations, Locations, and Disciplines., *J. Biomol. Screen.*, 2014, [Epub ahead of print]
- E. G. Brown, L. Wood and S. Wood, The medical dictionary for regulatory activities (MedDRA), *Drug Saf.*, 1999, **20**, 109-117.

- Anon. Elsevier, <u>www.elsevier.com/online-tools/pharmapendium (accessed: 03. 05.</u>
 2014).
- Anon. Thomson Reuters, <u>http://thomsonreuters.com/integrity/ (accessed: 03. 15.</u>
 2014)
- 37. M. A. van Driel, J. Bruggeman, G. Vriend, H. G. Brunner and J. A. Leunissen, A text-mining analysis of the human phenome., Eur. J. Hum. Genet., 2006, **14**, 535-542.
- 38. Anon. Tox21 Home Page

http://www.ncats.nih.gov/research/reengineering/tox21/tox21.html

- 39. Anon. ToxCast Initiative home page: <u>http://www.epa.gov/ncct/toxcast/</u>
- 40. Anon. Access to ToxCast data: http://epa.gov/ncct/toxcast/data.html
- 41. L. Yang, H. Luo, J. Chen, Q. Xing and L. He, SePreSA: a server for the prediction of populations susceptible to serious adverse drug reactions implementing the methodology of a chemical-protein interactome, *Nucleic Acids Res.*, 2009, **37**, W406-412.
- Wallach, N. Jaitly and R. Lilien, A structure-based approach for mapping adverse drug reactions to the perturbation of underlying biological pathways. *PLoS One*. 2010, **5**, e12063.
- 43. M. Keiser, V. Setola, J. Irwin, C. Laggner, A. I. Abbas, S. J. Hufeisen, N. H. Jensen,
 M. B. Kuijer, R. C. Matos, T. B. Tran, R. Whaley, R. A. Glennon, J. Hert, K. L.
 Thomas, D. D. Edwards, B. K. Shoichet and B. L. Roth Predicting new molecular targets for known drugs, *Nature*, 2009, **462**, 175-181.

- M. J. Keiser, B. L. Roth, B.N. Armbruster, P. Ernsberger, J. J. Irwin and B. K. Shoichet, Relating protein pharmacology by ligand chemistry, *Nature Biotechnol.*, 2007, 25, 197-206.
- 45. J. Hert, M. J. Keiser, J. J. Irwin, T. I. Oprea and B. K. Shoichet, Quantifying the relationships among drug classes, *J. Chem. Inf. Model.*, 2008, **48**, 755–765.
- 46. M. Glick, A. E. Klon, P. Acklin and J. W. Davies Enrichment of extremely noisy high-throughput screening data using a naïve Bayes classifier, J. Biomol. Screen., 2004, 9, 32-36.
- 47. K. Wang, J. Sun, S. Zhou, C. Wan, S. Qin, C. Li, L. He and L. Yang, Prediction of drug-target interactions for drug repositioning only based on genomic expression similarity, *PLoS Comput, Biol.* 2013, **9**, e1003315.
- 48. S. Zhao, T. Nishimura, Y. Chen, E. U. Azeloglu, O. Gottesman, C. Giannarelli, M. U. Zafar, L. Benard, J. J. Badimon, R. J. Hajjar, J. Goldfarb and R. Iyengar, Systems pharmacology of adverse event mitigation by drug combinations, *Sci. Transl. Med.*, 2014, **51**, 91-101.
- K. Briggs, M. Cases, D. J. Heard, M. Pastor, F. Pognan, F. Sanz, C. H. Schwab, T. Steger-Hartmann, A. Sutter, D. K. Watson and J. D. Wichard, Inroads to Predict *in Vivo* Toxicology-An Introduction to the eTOX Project, *Int. J. Mol. Sci.*, 2012, 13, 3820-3846.
- 50. J. L. Jenkins, A. Bender and J. W. Davies, *In silico* target fishing: Predicting biological targets from chemical structure, *Drug Discov. Today*, 2006, 3, 413–421.

- 51. T. M. Steindl, D. Schuster, C. Laggner, K. Chuang, R. D. Hoffmann and T. Langer, Parallel screening and activity profiling with HIV protease inhibitor pharmacophore models, *J. Chem. Inf. Model.*, 2007, **47**, 563–571.
- 52. T. I. Oprea, A. Tropsha, J. L. Faulon, M. D. Rintoul, Systems chemical biology, *Nat. Chem. Biol.*, 2007, **3**, 447–450.
- Anon. GVK-BIO. GVK Bio database suite. <u>http://www.gvkbio.com</u> (accessedJanuary 15, 2008).
- 54. Anon. Sunset Molecular Discovery LLC. WOMBAT: World of Molecular BioactiVity. 2007, <u>http://www.sunsetmolecular.com</u> (accessed January15, 2008).
- 55. Anon. Symyx Technologies. MDL Drug Data Report. <u>http://www.symyx.com</u> (accessed January 15, 2008).
- 56. S. Ekins, A. Bugrim, L. Brovold, E. Kirillov, Y. Nikolsky, E. Rakhmatulin, S. Sorokina, A. Ryabov, T. Serebryiskaya, A. Melnikov, J. Metz and T. Nikolskaya, Algorithms for network analysis in systems-ADME/Tox using the MetaCore and MetaDrug platforms, *Xenobiotica*, 2006, **36**, 877–901.
- 57. E. L. Berg, E. J. Kunkel, E. Hytopoulos and I. J. Plavec, Characterization of compound mechanisms and secondary activities by BioMAP analysis, *Pharmacol. Toxicol. Methods*, 2006, **53**, 67-74.
- 58. C. Luni, E. Serena and N. Elvassore, Human-on-chip for therapy development and fundamental science. *Curr. Opin. Biotechnol.*, 2014, **25**, 45-50.
- 59. M. B. Esch, A. S. Smith, J. M. Prot, C. Oleaga, J. J. Hickman and M. L. Shuler, How multi-organ microdevices can help foster drug development. *Adv. Drug Deliv. Rev.*, 2014, pii:S0169-409X(13)00289-5.

- I. Plavec, O. Sirenko, S. Privat, Y. Wang, M. Dajee, J. Melrose, et al., Method for analyzing signaling networks in complex cellular systems. *Proc. Nat. Acad. Sci.* USA, 2004, **101**, 1223–1228.
- J. D. Storey and R. Tibshirani, Statistical significance for genomewide studies. *Proc. Nat. Acad. Sci. USA*, 2003, **100**, 9440–9445.
- N. C. Kleinstreuer, J. Yang, E. L. Berg, T. B. Knudsen, A. M. Richard, M. T. Martin, D. M. Reif, R. S. Judson, M. Polokoff, D. J. Dix, R. J. Kavlock, K. A. Houck, Phenotypic screening of the ToxCast chemical library to classify toxic and therapeutic mechanisms. *Nat. Biotechnol.*, 2014, **32**, 583-591.
- M. A. Yildirim, K. I. Goh, M. E. Cusick, A. L. Barabási and M. Vidal, Drug-target network, *Nat. Biotechnol.*, 2007, 25, 1119-1126.
- 64. A. Cami, S. Manzi, A. Arnold, B. Y. Reis, Pharmacointeraction network models predict unknown drug-drug interactions. *Sci. Transl. Med.*, 2011, **3**, 114-127.
- 65. A. L. Hopkins, Network pharmacology: the next paradigmin drug discovery, *Nature Chem. Biol.*, 2008, **4**, 682–690.
- M. Kuhn, M. Campillos, I. Letunic, L. J. Jensen and P. Bork, A side effect resource to capture phenotypic effects of drugs, *Mol. Syst. Biol.*, 2010, 6, 343.
- S. Lee, K. H. Lee, M. Song and D. Lee, Building the process-drug-side effect network to discover the relationship between biological processes and side effects, *BMC Bioinforma*, 2011, **12**, S2.
- 68. J. Wang, Z. X. Li, C. X. Qiu, D. Wang, and Q. H. Cui, The relationship between rational drug design and drug side effects, *Brief Bioinform.*, 2012, **13**, 377-382.
- 69. S. Ekins, Progress in computational toxicology, *J. Pharm. and Tox. Methods*, 2013,69, 115-140.

- 70. A. Kinsner-Ovaskainen, G. Maxwell, J. Kreysa, J. Barroso, E. Adriaens, N. Alépée, N. Berg, S. Bremer, S. Coecke, J.Z. Comenges, R. Corvi, S. Casati, G. Dal Negro, M. Marrec-Fairley, C. Griesinger, M. Halder, E. Heisler, D. Hirmann, A. Kleensang, A. Kopp-Schneider, S. Lapenna, S. Munn, P. Prieto, L. Schechtman, T. Schultz, J. M. Vidal, A. Worth and V. Zuang, Report of the EPAA-ECVAM workshop on the validation of Integrated Testing Strategies (ITS), *Altern. Lab. Anim.*, 2012, **40**, 175-181.
- 71. A. Cherkasov, E.N. Muratov, D. Fourches, A. Varnek, I.I. Baskin, M. Cronin, J. Dearden, P. Gramatica, Y.C. Martin, R. Todeschini, V. Consonni, V. E. Kuzmin, R. Cramer, R. Benigni, C. Yang, J. Rathman, L. Terfloth, J. Gasteiger, A. Richard and A. Tropsha, QSAR Modeling: Where Have You Been? Where Are You Going To?, *J. Med. Chem.*, 2014, [Epub ahead of print].

Figure 1. Reverse translation used for identifying off-targets associated with ADRs.

A. Drugs with unrelated therapeutic targets and effects may produce common adverse reactions (ADRs) in patients. Association of the observed ADR with a common target by *in vitro* profiling of the drugs will enable the calculation of a therapeutic index (IT) based on available pharmacokinetic information. Most of the time, this is defined by the effective free maximum plasma concentration (EFPC). Once this process is complete and the effect of the drugs at the common target is confirmed by the TI, a common application is implemented.

B. *In vitro* target-profile of compounds in the safety pharmacology panel. Examples demonstrate the identification of a common off-target for Pergolide (anti-Parkinson drug) and nor-fenfluramine (metabolite of fenfluramine, a component used for the treatment of obesity). Both of these drugs cause cardiac valvular disease, a rare ADR, not observed by other drugs from the same class, unless 5-HT2B agonism is detected. Another anti-Parkinson drug, Ropinirole, and Rimonabant, a CB1 antagonist drug for the indication of obesity do not show 5-HT2B agonism and consequently do not cause cardiac valvular disease. Red arrows point to results obtained in the 5-HT2B assay (agonist mode).

Figure 2. A. The diagram demonstrates the application of the ligand-based in silico approach for off- target identification using reverse translation. Ligand-target-ADR networks can be generated and searched for unknown off-targets or in a more complex model pathways responsible for adverse phenotypes. B. Practical application of the method shows the discovery of the inhibitory effect of chlorotrianisene (a synthetic

estrogen) at the cyclooxygenase-1 (COX-1) enzyme. In addition to the shared ADRs (erythema multiforme and edema) associated with both estrogen receptor activation and COX-1 inhibition, upper abdominal pain and rush predicted that COX-1 inhibition, a so far unknown off-target was inhibited by chlorotrianisene. (reprinted with permission from Lounkine et al.²¹)

Figure 3. Example of computational workflow to identify compound features associated with target-based toxicities within an adverse event (AE) network. Compounds sharing common toxicity are extracted and targets for each compound are predicted by using a multiple category Bayes model capable of predicting over 2000 targets. The predicted targets are then put into the context of pathways and generate a target-ADR network. This way toxicity can be linked to the pathways responsible for the undesired effects and generate an integrated molecular network.





В



Figure 2.



В



Toxicology Research Accepted Manuscrip

Figure 3.

