In silico models in drug development: where we are

Janet Piñero*, Laura I. Furlong*, Ferran Sanz**

Research Programme on Biomedical Informatics (GRIB), Department of Experimental and Health Sciences (DCEXS), Hospital del Mar Medical Research Institute (IMIM), Universitat Pompeu Fabra (UPF), Carrer Dr. Aiguader 88, 08003 Barcelona, Spain

* These authors contributed equally

** Corresponding author: ferran.sanz@upf.edu

Abstract

The use and utility of computational models in drug development has significantly grown in the last decades, fostered by the availability of high throughput datasets and new data analysis strategies. These *in silico* approaches are demonstrating their ability to generate reliable predictions as well as new knowledge on the mode of action of drugs and the mechanisms underlying their side effects, altogether helping to reduce the costs of drug development. The aim of this review is to provide a panorama of developments in the field in the last two years.

Highlights

- A variety of computational methods and tools are used in QSP and QST, with different degrees of maturity.
- PBPK models are well established and applied in a wide variety of scenarios.
- Pharmacogenomics and toxicogenomics data are employed to gain mechanistic understanding
- Cellular signaling models are mostly used to predict treatment response in cancer.
- GSMN models allow easy integration of mechanistic knowledge.
- Multiscale and multi-component models constitute the frontier of the field.

Introduction

Quantitative Systems Pharmacology (QSP) is a relatively new discipline that combines systems biology approaches with methods of quantitative pharmacology [1]. The combination of

computational and experimental methods via QSP approaches provides a systems level understanding of the mechanism of action of drugs while leveraging on the accumulated data on approved or failed drugs. In a similar way, Quantitative Systems Toxicology (QST), emerged as new paradigm for toxicity assessment [2], focuses on understanding the adverse effects of drugs, from molecular alterations to phenotypic observations, by integrating computational and experimental methods [3]. QST merges methods of classic toxicology with systems biology modeling and quantitative measurements of molecular and functional changes occurring upon drug treatment at different levels of biological organization (cell, tissue, organ, organism) [2]. QST approaches have proven useful to optimize dose and schedule drug regimens, potentially minimizing costly phase I/II clinical trials [4,5]. By integrating *in vitro* cell toxicity data with multiscale *in silico* modeling of drug exposure, QST models could become an efficient tool to assess and predict drug toxicity [3]. Moreover, a better understanding of biological responses to drugs will reduce uncertainties in species extrapolations, and allow the prediction of treatment responses considering the patient genetic variability or pre-existing diseases.

The present review is focused on presenting and discussing the recent advancements in computational methods used in QSP and QST, which support three crucial aspects of the drug development process: i) the understanding and prediction of drug pharmacokinetics, ii) the understanding and prediction of drug toxicity, and iii) the translational perspective of the pre-clinical assessment.

Physiologically based pharmacokinetic models

Physiologically based pharmacokinetic and pharmacodynamic (PBPK/PD) modeling has become a widely adopted tool in the industry to obtain a quantitative characterization of concentration–time profiles in different organ and tissues across human populations. A recent survey showed that around 70% of pharmaceutical companies use pre-clinical PBPK/PD modeling in all therapeutic areas [6]. The wide adoption of these modeling approaches has been facilitated by the availability of several PBPK commercial platforms [7], and by recommendation of regulatory agencies [8]. The main goal of PBPK modeling is to describe drug absorption, distribution, metabolism and elimination (ADME) within the body. The prediction of drug exposure in plasma but especially in the site of action of the drug is of high pharmacological relevance, because drug concentration in certain body compartments may be

difficult or impossible to be experimentally measured [7]. State of the art PBPK/PD models are composed of hundreds of ordinary differential equations (ODEs) describing physiological processes involved in ADME. The parameters in the model are obtained from prior knowledge available in the literature or calculated from specific and carefully validated formulas [7]. Although the primary focus of a PBPK model is on physiological variables, biochemical information is considered for drug transporters and metabolic enzymes, which play a role in drug transport and metabolism.

PBPK models have been used to represent particular disease states or specific patient groups, such as pediatric patients or pregnant women [9] as well as to predict drug-drug interactions [10–16], food-drug interactions [17–19], drug formulation effects [20,21], cross-species extrapolation [22–24], and constitute key components of multiscale models [25]*.

PBPK models can be combined with transcriptomics data to investigate mechanisms of drug toxicity [26,27]* and carcinogenicity [28]. Furthermore, PBPK models can be expanded by adding mechanistic models of gene regulation and signaling pathways. For instance, a PBPK model was coupled with the miRNA-BDNF pathway to study perfluorooctanesulfonic acid induced neurotoxicity [29]. In another study, Mason et al. combined PK and mechanistic models to estimate the dose and time of ingestion in paracetamol poisoning, using traditional and experimental serum biomarkers in mice [30]**.

Although PBPK models are widely used for the prediction of ADME, other types of modeling approaches are required to gain insight on the mode of action of compounds, especially at the cellular level.

Toxicogenomics data analysis

The use of transcriptomics to characterize the cell response to a particular compound is widely applied in both QSP and QST. DNA microarray technologies have allowed monitoring the changes of the expression levels of thousands of genes simultaneously after the exposure to a given drug, setting the foundations for the field of toxicogenomics. The most popular resources for toxicogenomics are summarized in Box 1. One of the challenges in the field is how to translate changes in gene expression into actionable information for understanding the biological mechanism of toxicity of drugs. To address this challenge, several approaches have been

proposed, including the analysis of gene signatures, gene set enrichment analysis, and gene co-expression networks.

Gene signatures analysis

Gene signatures analysis aims at obtaining a minimal list of genes that can be used to predict the toxic response to a compound. The underlying assumption is that compounds with similar mechanisms of action will have similar gene expression profiles, and that these gene expression profiles can be used to build gene expression signatures predictive of drug toxicity. A variety of methodologies have been proposed to identify these gene signatures. Among them, Connectivity Map-like analysis [31] aims at detecting similarities among gene expression signatures of different compounds using pattern-matching algorithms. This method has been successfully used to group chemicals based on their mode of action [32], to select potential new drug candidates for several cancer types [33], to characterize genes involved in the cell response to different chemicals by means of different features, such as evolution, topological properties in a protein interaction network and disease SNP density [34], and by integrative analysis with chemical structures and drug sensitivity data, to improve drug taxonomy and provide a comprehensive picture of drug-drug relationships [35]*.

Another type of methods uses machine–learning techniques to derive the gene signatures. For example, Rempel *et al.* obtained a classifier that allows to separate histone deacetylase inhibitors from mercurials using human embryonic stem cells, thus demonstrating that the system is suitable for toxicant classification [36] and Giordano et al used different machine-learning approaches to derive gene signatures from whole blood gene expression data to predict cigarette smoke exposure in humans [37].

Recently, a crowdsourcing-based project annotated and re-analyzed different types of gene expression profiles from Gene Expression Omnibus (GEO), including approximately 1,000 drug perturbation signatures [38]. The manually curated signatures were used as training set to develop classifiers for extracting similar signatures from the entire GEO repository, and were made available at the Crowd Extracted Expression of Differential Signatures (CREEDS) web portal. Finally, TOXsIgN (for TOXicogenomic sIgNatures) is a resource that supports the online submission, storage and retrieval of toxicogenomic signatures of hundreds of compounds in humans, rats, mice or drosophila [39].

Gene set enrichment analysis

Other way to analyze toxicogenomics data uses functional gene set enrichments to reduce dimensionality of the data and the experimental noise, and to suggest plausible biological hypotheses that explain the cellular response to drug treatment. The functional gene sets are obtained from resources such as Reactome [40], Wikipathways [41], Gene Ontology [42] or MsigDB [43]. Traditionally, the differentially expressed genes are compared to the gene sets to identify significant overlaps. This type of approach was used by Parmentier et al to characterize the gene signature underlying cholestasis in a modified human hepatocyte 2D-sandwich culture treated with five drugs [44].

A very popular method for analyzing gene expression data is gene set enrichment analysis (GSEA) [45]. In GSEA, genes are ranked based on a certain metric, for example, the expression level, with the goal to determine whether members of a given gene set tend to be located at the top (or the bottom) of the ranked gene list, in which case the gene set is assumed to be correlated with the associated phenotype or condition measured in the experiment. In this way, modest but coordinated changes in predefined sets of genes can be detected. In the context of toxicogenomics, GSEA has been able to pinpoint enriched pathways that could inform of the mode of action of a drug in a dose- and time-dependent manner. By combining this method with benchmark dose modeling [46], it is possible to estimate doses at which different cellular pathways are altered in toxicogenomic experiments [47]. GSEA combined with structural data was used to detect compounds with similar structure that induce different transcriptional responses, and vice versa, drugs that elicit similar transcriptional responses but differ in their chemical structure [48]. In another example, GSEA analysis combined with machine learning was used to build a predictive toxicogenomics space (PTGS) tool, composed of over 1,300 genes distributed over 14 overlapping cytotoxicity-related gene space components. The tool is able to predict dose-dependent liver toxicity in hepatocytes, and drug-induced liver injury (DILI) in humans [49]**.

Another method, Dose-Time Network Identification (DTNI) [50]*, infers gene co-expression networks from toxicogenomics data considering both dose and timing of drug exposure. The method is based on a system of Ordinary Differential Equations (ODEs), whose parameters are estimated using regression techniques from the data and represent the interaction

strengths between genes. DTNI allows finding compounds that affect specific cellular pathways, and infering new gene interactions. DTNI was applied to reconstruct gene regulatory networks of four stress-related pathways (TP53, ER, NRF2, and NF-kB) in order to infer causal evidence for mechanisms explaining DILI and carcinogenicity [51]. In another recent example, Gene Ontology enrichment analysis of the gene signatures of 33 compounds in TG-GATEs was integrated with biomedical literature mining to describe human diseases associated to the same compounds and to identify links that support the suitability of *in vitro* and *in vivo* systems to model the physiological effects of drugs on humans [52].

Gene co-expression networks

Gene set enrichment analysis is biased toward known biology captured in existing collections of gene sets and pathways. This is why co-expression network analysis, which does not rely on previous biological knowledge of the system, has emerged as an alternative data-driven, unsupervised approach, which uses the property of gene co-expression upon drug perturbation to organize genes into networks.

The gene co-expression modules can be generated using different methodologies. For example, gene co-expression modules obtained with the Iterative Signature Algorithm [53] were used to predict acute kidney [54] and liver injuries [55]. While most studies relies on mRNA gene expression in response to drugs, Pang et al focused on regulation of gene expression by non-coding RNA to develop gene co-expression networks between lncRNAs and mRNAs, to pinpoint lncRNAs that could act as biomarkers of bisphenol A-induced neurotoxicity [56]. Another approach to obtain clusters from co-expression networks is weighted gene co-expression network analysis (WGCNA) [57]. Using this methodology, the TXG-MAP ('toxicogenomic module associations with pathogenesis') approach was proposed to characterize mechanisms of DILI [58]**. WGCNA approaches have also been used for drug repositioning in cancer using gene expression datasets from nine major human cancer types [59].

A recent survey carried out in the pharmaceutical industry has shown that toxicogenomic analyses are performed early in drug discovery with the aim to gather mechanistic insight on drug mode of action, but they are not yet widely used to predict the toxic effects of drugs [60]. One reason that explains the paucity in the adoption of this approach in toxicity assessment is the difficulty in the interpretation of the results [61]. First, it is challenging to distinguish

changes in gene expression due to the physiological response to the drug from the ones related to toxicity mechanisms. Second, not all the drug toxicity effects will be related to changes in gene expression profiles. Third, since multiple cellular signaling pathways may converge to alter the expression of the same gene products, the identification of the upstream pathway responsible for the gene expression changes upon drug perturbation is not a trivial task. Finally, most drugs act through multiple mechanisms of action that depend on dose, timing and duration of exposure, and the particular condition and phenotype of the cell in which they act [61].

Modeling cellular signaling networks

Systems biology modeling can help to understand the mode of action of drugs and predict the behavior of a biological system in response to drug perturbations [62]. The molecular pathways used by cells to interpret signals from the environment are not linear, but interconnected by cross-talk mechanisms, giving rise to complex signaling networks. The action of drugs on these intricate cellular networks can result on either therapeutic or toxic effects [62]. Although a variety of systems biology approaches can been used in QSP and QST, the choice of the modeling approach to represent the cellular processes underlying drug action depends on the number of components to model, the type of data available, and the biological questions under investigation. Some of the most commonly used modeling strategies include physicochemical models based on sets of coupled ordinary differential equations (ODEs), logic modeling and graph-based approaches.

ODE-based models are mechanistic, dynamical models that describe the behavior of the system over time using mass-action kinetics for the rates of production and consumption of the molecular species. A mature field in terms of dynamic, mechanistic models is cardiac electrophysiology, where ODE models are available since 1960 [63]. The O'Hara-Rudy dynamic (ORd) cardiac ventricular model, consisting of 4 compartments, 15 ion channels, 6 ionic fluxes, 5 buffers, and CaMK [64] is one of the most widely used cardiac model, and has been recently adopted by regulatory bodies to base decisions on cardiovascular safety [65]. The ORd model has been extended in several ways to broaden its applicability, for instance to model gender differences in risk of arrhythmia [66], or by considering dynamic drug–hERG interactions and multichannel pharmacology [67]. Recently, Passini *et al.* showed that models of human ventricles based on the ORd model could better predict clinical risk of Torsade de Pointes

arrhythmias than experiments on animal models or on human induced pluripotent stem cell-derived cardiomyocytes [68]**. Similarly, an *in silico* tool has been developed to predict drug induced alteration in the action potential and the QT interval by means of a multiscale model that incorporates molecular simulations of the blocking of several channels and simulations of the electrophysiology of a virtual tissue using ORd models, and that takes into account the effective free therapeutic plasma concentration of the drug [69]**.

Although quantitative, mechanistic models such as ODE-based are preferred, we often lack the detailed knowledge on biochemical processes required to apply them [70]. Therefore, logic models are used to describe processes comprising medium or large scale networks where detailed biological knowledge is incomplete [71]. This type of models has been applied to cancer, due to the relevance of signaling and regulatory networks in cancer development and therapy and the availability of relevant omics data. Different formalisms can be used to model signaling and regulatory networks (for a review see [72]). For instance, Boolean dynamic models were able to predict resistance to PI3K inhibitors in breast cancer and to suggest novel combinatorial therapies more effective than PI3K inhibition alone [73]. Stochastic Boolean network models were used to pinpoint candidate genes that, if suppressed, might have an impact in cell viability in breast cancer [74]. Synchronous and asynchronous Boolean networks can be used to predict therapeutic targets, for example, in bladder tumorigenesis [75]. Boolean models derived from signaling pathways and gene expression data were combined with drug target information and protein interaction data to prioritize candidates for drug repurposing in Triple Negative Breast Cancer, a subtype of breast cancer without specific therapeutic targets [76].

Morris *et al.* used a constrained fuzzy logic model of the signaling network activities in hepatocellular carcinoma to predict drug combinations of kinase inhibitors for the treatment of this cancer type [77]. The novelty of their approach resides in simulating the microenvironment of the tumor that modulates drug response by taking into account growth factors and inflammatory cytokines. This model predicted successfully drug combinations inhibiting tumor-promoting signaling activities under diverse stimulation conditions.

Finally, graph-based methods are qualitative modeling approaches based on the structure of the network, which do not require information on kinetic parameters, and are therefore applicable to large-scale networks [78], in contrast to ODE-based approaches. Typically, they enable the identification of feedback loops and signal transfer routes in signalling networks. An

example of such approach was recently implemented using the Integer Linear Programming (ILP) algorithm to identify the mode of action of drugs [79]*. More specifically, for a given drug, the ILP identifies paths starting from the targets of the drug, through the signaling cascade to the transcription factors and finally arriving to the gene expression level. ILP was used to identify a cellular signaling network underlying drug induced lung injury using gene expression data for over 200 drugs from the Connectivity Map in combination with a knowledge-based functional network including protein interaction data and transcription factors information. The method allowed not only establishing the mode of action for the drugs and proposing molecular mechanisms underlying lung injury, but also predicting candidate drugs to treat lung injury [79]. A similar approach was used to build predictive models for drug-induced cardiomyopathy [80].

Genome scale metabolic modeling

A key aspect of drug response is how drugs are metabolized in the body. Thus, multi-scale QSP and QST models often include a component to model drug metabolism. In this regard, the development of Genome Scale Metabolic Networks (GSMNs) has allowed the mechanistic modeling of human metabolism at the level of cells, tissues and organs [81]. GSMNs are composed of hundreds of ODEs that are solved by constraint-based modeling (e.g. Flux Balance Analysis). To construct a GSMN, different types of data are integrated, including information of all known biochemical and transport reactions extracted from the literature, which are complemented with genomics, proteomics, and metabolomics data. Several human GSMN are currently available, including a global [82], and tissue-specific GSMNs, such as kidney [83], as well as cell-specific models for hepatocytes [81,84], adipocytes [85], enterocytes [86], and myocytes [87]. GSMNs have also been constructed for several organisms using a variety of tools [88]. A comprehensive GSMN for rat has been recently published [89]**. This rat GSMN has been used in combination with gene expression data to predict metabolites sensitive to particular toxicity response for rat and human hepatocytes exposed to more than 70 environmental toxicants and pharmaceuticals. To achieve this goal, they developed an algorithm called TIMBR that is able to estimate how the production of a metabolite is affected by changes in gene expression [89].

GSMNs have also been used to predict chemical-induced hepatotoxicity, by integrating toxicity assays and transcriptomic data. This multilevel integration also allowed to characterize

the main metabolic pathways altered upon chemical perturbations which in turn, provide insight into the mechanisms of toxicity [90]. A similar approach was used to model drug-drug interactions between phenytoin and oral estradiol [91]. GSMN models can be combined with information on drug targets and drug side-effects to build models that allow to predict if a drug will induce a side-effect by examining the drug's impact on genome-wide metabolic fluxes [92].

Multiscale modeling

Computational models that integrate different spatial, temporal and functional scales, to enable the description and simulation of the emergent properties of a system are now commonly applied in several areas, including ecology and human disease [93]. The development of the field has been boosted by the advent of omics technologies, capable of producing systems-level measurements for multiple types of biomolecules. Drug development has been proposed as "the ultimate multiscale optimization problem" since drug response is examined across temporal and spatial scales along all the phases spanned by the process of drug discovery: cell cultures, tissues, organs, organisms, and, finally, human populations [94]. Not surprisingly, many strategies of multiscale modeling in QSP and QST incorporate PBPK models because, on one hand, PBPK models are able to simulate effective doses across several levels of organization (cell, tissue, organ) and, on the other, they are mature and routinely used in drug development. A multiscale model for liver xenobiotic metabolism and toxicity that incorporates whole body, tissue (hepatic sinusoid), and sub-cellular levels (pathways for Phase I and Phase II metabolism) PBPK was developed by Sluka, *et al.* [95]. As a case study, they showed the results of evaluating the model for pharmacological doses of acetaminophen.

Maldonado *et al.* [25]** proposed a methodology to expand PBPK models to incorporate whole-cell metabolism and gene expression regulation of key drug metabolizing enzymes. This approach introduces dependencies between the different model components combined in the multiscale model. The model uses PBPK, Hepatonet1 GSMN and the gene regulatory network of CYP3A4 developed with the MUFINS framework to describe the pharmacokinetics of paracetamol and its toxic metabolite NAPBQI in the context of patient chronic stress and patient-specific liver metabolism [25]. This model allows for instance to identify metabolic reactions in the liver that can alter the production of GSH, and therefore affect the detoxification of drug metabolites and xenobiotics. Although the model was developed for illustrative purposes and not

to model the particular compound, the approach can be used to identify the enzymes catalyzing these reactions as potential pharmacogenes of interest for the toxicity of drugs.

The integration of GSMN models with PBPK models can expand the later to account for whole cell metabolism, linking them to genomic information about metabolic enzymes, and enabling mechanistic assessment of drug-induced metabolic perturbations [25]. For instance, a GSMN model of a human liver was combined with a whole-body PBPK model for isoniazid to explain the mechanisms underlying DILI of this compound [96]. This combined model quantitatively shows how the isoniazid-induced metabolic perturbations are distributed and attenuated in the liver. Furthermore, the predictions of the model for several metabolites related to liver physiology, such as cholesterol, amino acids, and fatty acids were found to be in agreement with results in patients and animal models.

Likewise, multiscale models have been developed to characterize the effects of drugs in cardiac electrophysiology [69,97]. Sahli Costabal *et al.* [97] modeled the propagation of the action potential duration at the cellular level, the excitation pattern across the left and right ventricles, and the QT interval at the organ level. To test the model, they used drugs that differed in the risk of producing Torsades de Pointes to show how the electrophysiological abnormalities propagate, from specific channel blockage, via altered single cell action potentials and prolonged QT intervals, to the emergence of ventricular tachycardia.

Translation from pre-clinical models to the clinical scenario

A deeper understanding of the differences in the physiological responses upon drug perturbation across organisms, but also across different type of experimental models (e.g. cell culture, organoids), should help to reduce uncertainties in model extrapolation [98]. Towards this goal, results from comparative genomics analysis between mouse and human show a global conservation of gene expression profiles, although the degree of conservation varies depending on the tissues and the genes that are compared [99]. In particular, a recent large scale study performed a characterization of the response to more than 100 different chemicals in rat liver *in vivo* and rat and human primary hepatocytes *in vitro*, using a modified GSEA [100]*. This analysis showed that the early toxicological response *in vivo* is recapitulated in human and rat primary hepatocyte cultures at the molecular level, indicating that these models are concordant in identifying key pathways in response to chemical stress. Strikingly, these results contradict those

of Sutherland *et al.*, who performed a study using data from Drug Matrix and TG-GATEs and different models (rat and mouse liver and human and rat primary hepatocytes), and found high concordance for the same model and source of the data, but low concordance in the response in rodent liver versus cultured hepatocytes [101]. Importantly, these results show marked transcriptional changes induced by cell culture, comparable in magnitude to highly toxic drug treatments to rat liver. The main reason behind this seemingly contradictory results may lie in the way they evaluated the cellular response: while El-Hachem *et al.* [102] evaluated only a small number of pathways to predict toxicity, Sutherland *et al.* assessed the effects of the perturbations at genome scale.

On the other hand, Shankaran *et al.* [103]* addressed the quantitative translation of predicted pre-clinical gastrointestinal (GI) toxicity of oncologic agents to the clinical scenario. The model was developed distinguishing the "system-specific" parameters governing GI physiology from the "drug-specific" parameters leading to drug toxicity. The system parameters, obtained from the literature, enabled to capture differences in GI physiology between rodents and humans. Their mathematical model incorporates known biology to predict GI toxicity and optimize dosing schedules for irinotecan to minimize clinical toxicity, based on rat GI toxicity data. The model is based on ODEs of intestinal cell dynamics that account for species-specific differences in GI turnover. The model was fit with pre-clinical data in irinotecan effects in rat and predicted human GI toxicity kinetics. It also succeeded to predict GI toxicity for novel dosing schedules, not yet tested in patients.

If we consider the metabolic point of view, there seems to be more translational concordance: the comparison between the reconciled rat and human metabolic networks showed that they differed only in eight enzymes, out of more than 2,000 proteins [89].

Incorporation of human variation data into QST and QSP models

Considering that inter-individual variation in drug response is essential to achieve safer drug treatment, and a necessary step towards personalized medicine, QSP and QST models should incorporate information on human genomic variation. In order to account for the genetic variation in human patients, a model has to include mechanistic details at the molecular level. Mechanistic models such as PBPK, GSMN, and especially systems biology models are well suited for assimilating this type of information. For example, Mih *et al.* integrated protein

structural information into GSMN models to explore the effects of sequence variation on drug responses in human erythrocyte metabolism [104]**. The assessment of the impact of genetic polymorphisms in genes involved in drug metabolism contained in PBPK models can help, for example, to identify individuals with variants leading to drug concentration outside the therapeutic window [25], or with an increased susceptibility to a certain type of toxicological event, for example with higher susceptibility to depolarization abnormalities. Nevertheless, in this type of approach, the consideration of the genetic variation would be restricted to genes included in the models, which constitute still a very small number (current PBPK models include approximately 20 genes involved in drug metabolism [25]). By coupling GSMN, signaling network modeling and other types of systems biology models with PBPK models, the scope of genomic variability in the genes modeled can be expanded by using data from large scale genomic initiatives such as EXAC [105], 1000 Genomes [106] and Genome UK [107], making it possible to predict the effect of genomic variation on drug response in specific populations.

Conclusion and future perspectives

Although both, QSP and QST have experienced tremendous advances in the last few years, there is still a need to develop fit-for-purpose, mechanistic, quantitative, multiscale models to improve toxicity assessment. In this regard, recent initiatives such as TransQST will pave the way for the development of multi-scale, quantitative QST models for cardiac, liver, kidney and gastrointestinal drug toxicity (http://transqst.org/). To foster the development of more accurate mechanistic models, the data gaps concerning the information on regulatory networks that take place in specific cell types and particular conditions should be addressed. The availability of detailed catalogues of proteins such as the human plasma proteome [108] is a step in this direction. On the other hand, the efforts that are being made for collecting and sharing great amount of high quality, toxicology-related data from the archives of the pharmaceutical companies through the eTOX (http://www.etoxproject.eu/) and eTRANSAFE (http://etransafe.eu/) projects, will also support the development more reliable predictive models [109]**.

It is important to bear in mind that the initial pathophysiological state of the biological system is defined by the interplay among genetic, epigenetic and environmental factors, which then dynamically evolve when perturbed by a drug. It has been shown that ADME genes can be

subjected to epigenetic modifications in a variety of diseases such as several cancer types, Parkinson disease and conditions such as smoking (for a review, see [110]). On the other hand, exposure to chemicals and drugs can alter the epigenetic status of an individual. Moreover, "epidrugs", or drugs that target the epigenetic machinery of the cell are already starting to be used for cancer treatment. The emerging fields of pharmacoepigenetics and toxicoepigenetics will help to understand the complex interplay between drugs and environmental factors, which will in turn, provide a deeper understanding of drug response.

Another important factor to be considered in QSP and QST models is the human microbiota. The role played by the human microbiota in the response to drug treatment has been increasingly recognized. Human microbiota is known to transform drugs into metabolites with pharmacological properties that could be toxic, teratogenic, and even lethal [111]. On the other hand, some drugs rely on the microbiota to be converted from inactive precursors to pharmaceutically active compounds [112]. Recently, genome-scale metabolic reconstructions of the human gut microbiota were generated [113]. These reconstructions can be used to produce hybrid PBPK – GSMN models to understand drug-microbiota-diet interactions [114]. This type of analysis will pave the way for the simulation of personalized microbiomes in QSP and QST models, and thus enabling a more comprehensive and precise simulation of drug effects.

Finally, it has to be pointed out that QSP and QST models are being accepted as key components of the drug development process. As a proof of this, pharmaceutical companies have incorporate them into their R&D pipeline and are funding related projects in the framework of the European Innovative Medicines Initiative, such as eTOX (http://www.etoxproject.eu/), eTRANSAFE (http://etransafe.eu/) and, particularly, TransQST (http://transqst.org/). At the same type, international bodies, such as the Organisation for Economic Co-operation and Development (OECD), published guidelines for the validation of computational models to be used in production environments [115], in this way demonstrating their importance and maturity.

Acknowledgements

We acknowledge support from ISCIII-FEDER (CPII16/00026), the EU H2020 Programme 2014-2020 under grant agreements no. 681002 (EU-ToxRisk) and no. 676559 (ELIXIR-EXCELERATE), and the IMI2-JU under grants agreements no. 116030 (TransQST) and no. 777365 (eTRANSAFE), resources of which are composed of financial contribution from the EU-H2020 and EFPIA companies in kind contribution. The Research Programme on Biomedical

Informatics (GRIB) is a member of the Spanish National Bioinformatics Institute (INB), funded by ISCIII and FEDER. The DCEXS is a "Unidad de Excelencia María de Maeztu", funded by the MINECO (ref: MDM-2014-0370).

Box 1: Glossary of terms or abbreviations.

ADME: drug absorption, distribution, metabolism and elimination
DILI: Drug-Induced Liver Injury
DTNI: Dose-Time Network Identification
ILP: Integer Linear Programming
GEO: Gene Expression Omnibus
GSEA: Gene Set Enrichment Analysis
GSMN: Genome Scale Metabolic Networks
ODE: Ordinary Differential Equations
ORd: O'Hara-Rudy dynamic cardiac ventricular model
PBPK/PD: Physiologically based pharmacokinetic/pharmacodynamics
QSP: Quantitative Systems Pharmacology
QST: Quantitative Systems Toxicology
WGCNA: Weighted Gene Co-expression Network Analysis

Box 2: Toxicogenomics data resources

One of the most commonly used resources in QSP and QST analysis is open access TG-GATEs database [116]. This resource contains toxicogenomics data for 170 compounds, in human and rat primary hepatocytes, linked to phenotype data and pathology findings. The US Broad Institute Connectivity Map [117,118] contains thousands of gene expression profiles of most FDA approved drugs tested in multiple cell types. It has been used for identifying modes of action and defining biologically similar compounds. The US National Cancer Institute (NCI) 60 tumor cell line screen includes results on GI50 (50% growth inhibition), total growth inhibition (TGI), and LC50 (50% lethal concentration) for many compounds tested in the major Connectivity Map cell lines [119]. The Library of Integrated Network-based Cellular Signatures (LINCS) catalogs how cells respond to different types of perturbations using a variety of assays [120]. The Chemical Effects in Biological Systems (CEBS) database is a toxicology resource containing animal data from the National Toxicology Program (NTP) testing program and other depositors. CEBS currently covers over 8,000 studies including carcinogenicity, short-term toxicity and genetic toxicity studies [121].

Bibliography

1. Sorger PK, Allerheiligen SRB, Altmann RB, Shoichet B, Lauffenburger D a, Iyengar R, Califano A: **Emergence of Quantitative and Systems Pharmacology: A White Paper**. *NIH QSP Work* 2011,
2. Sturla SJ, Boobis AR, FitzGerald RE, Hoeng J, Kavlock RJ, Schirmer K, Whelan M, Wilks MF, Peitsch MC: **Systems toxicology: from basic research to risk assessment.** *Chem Res Toxicol* 2014, **27**:314–29.
3. Bloomingdale P, Housand C, Apgar JF, Millard BL, Mager DE, Burke JM, Shah DK: **Quantitative systems toxicology.** *Curr Opin Toxicol* 2017, **4**:79–87.
4. Cadoo KA, Gajria D, Suh E, Patil S, Theodoulou M, Norton L, Hudis CA, Traina TA: **Decreased gastrointestinal toxicity associated with a novel capecitabine schedule (7 days on and 7 days off): a systematic review.** *NPJ breast cancer* 2016, **2**:16006.
5. Fang L, Xin W, Ding H, Zhang Y, Zhong L, Luo H, Li J, Yang Y, Huang P: **Pharmacokinetically guided algorithm of 5-fluorouracil dosing, a reliable strategy of precision chemotherapy for solid tumors: a meta-analysis**. *Sci Rep* 2016, **6**:25913.
6. Schuck E, Bohnert T, Chakravarty A, Damian-Iordache V, Gibson C, Hsu C-P, Heimbach T, Krishnatry AS, Liederer BM, Lin J, et al.: **Preclinical Pharmacokinetic/Pharmacodynamic Modeling and Simulation in the Pharmaceutical Industry: An IQ Consortium Survey Examining the Current Landscape**. *AAPS J* 2015, **17**:462–473.
7. Kuepfer L, Niederalt C, Wendl T, Schlender J-F, Willmann S, Lippert J, Block M, Eissing T, Teutonico D: **Applied Concepts in PBPK Modeling: How to Build a PBPK/PD Model**. *CPT Pharmacometrics Syst Pharmacol* 2016, **5**:516–531.
8. van der Graaf PH, Benson N: **Systems Pharmacology: Bridging Systems Biology and Pharmacokinetics-Pharmacodynamics (PKPD) in Drug Discovery and Development**. *Pharm Res* 2011, **28**:1460–1464.
9. Ke AB, Greupink R, Abduljalil K: **Drug Dosing in Pregnant Women: Challenges and Opportunities in Using Physiologically Based Pharmacokinetic Modeling and Simulations**. *CPT Pharmacometrics Syst Pharmacol* 2018, **7**:103–110.
10. Xue C, Zhang X, Cai W: **Prediction of Drug-Drug Interactions with Bupropion and**

Its Metabolites as CYP2D6 Inhibitors Using a Physiologically-Based Pharmacokinetic Model. *Pharmaceutics* 2017, **10**:1.

11. Moj D, Hanke N, Britz H, Frechen S, Kanacher T, Wendl T, Haefeli WE, Lehr T: **Clarithromycin, Midazolam, and Digoxin: Application of PBPK Modeling to Gain New Insights into Drug–Drug Interactions and Co-medication Regimens**. *AAPS J* 2017, **19**:298–312.
12. Pilla Reddy V, Walker M, Sharma P, Ballard P, Vishwanathan K: **Development, Verification, and Prediction of Osimertinib Drug-Drug Interactions Using PBPK Modeling Approach to Inform Drug Label**. *CPT Pharmacometrics Syst Pharmacol* 2018, doi:10.1002/psp4.12289.
13. Gu H, Dutreix C, Rebello S, Ouatas T, Wang L, Chun DY, Einolf HJ, He H: **Simultaneous Physiologically Based Pharmacokinetic (PBPK) Modeling of Parent and Active Metabolites to Investigate Complex CYP3A4 Drug-Drug Interaction Potential: A Case Example of Midostaurin.** *Drug Metab Dispos* 2018, **46**:109–121.
14. Marzolini C, Rajoli R, Battegay M, Elzi L, Back D, Siccardi M: **Physiologically Based Pharmacokinetic Modeling to Predict Drug–Drug Interactions with Efavirenz Involving Simultaneous Inducing and Inhibitory Effects on Cytochromes**. *Clin Pharmacokinet* 2017, **56**:409–420.
15. Duan P, Zhao P, Zhang L: **Physiologically Based Pharmacokinetic (PBPK) Modeling of Pitavastatin and Atorvastatin to Predict Drug-Drug Interactions (DDIs)**. *Eur J Drug Metab Pharmacokinet* 2017, **42**:689–705.
16. Marsousi N, Desmeules JA, Rudaz S, Daali Y: **Prediction of drug-drug interactions using physiologically-based pharmacokinetic models of CYP450 modulators included in Simcyp software**. *Biopharm Drug Dispos* 2018, **39**:3–17.
17. Li X, Shi L, Tang X, Wang Q, Zhou L, Song W, Feng Z, Ge J, Li JK, Yang L, et al.: **Mechanistic prediction of food effects for Compound A tablet using PBPK model**. *Saudi J Biol Sci* 2017, **24**:603–609.
18. Rose RH, Turner DB, Neuhoff S, Jamei M: **Incorporation of the Time-Varying Postprandial Increase in Splanchnic Blood Flow into a PBPK Model to Predict the Effect of Food on the Pharmacokinetics of Orally Administered High-Extraction Drugs**. *AAPS J* 2017, **19**:1205–1217.

19. Cristofoletti R, Patel N, Dressman JB: **Differences in Food Effects for 2 Weak Bases With Similar BCS Drug-Related Properties: What Is Happening in the Intestinal Lumen?** *J Pharm Sci* 2016, **105**:2712–2722.
20. Purohit HS, Trasi NS, Sun DD, Chow ECY, Wen H, Zhang X, Gao Y, Taylor LS: **Investigating the Impact of Drug Crystallinity in Amorphous Tacrolimus Capsules on Pharmacokinetics and Bioequivalence Using Discriminatory In Vitro Dissolution Testing and Physiologically Based Pharmacokinetic Modeling and Simulation**. *J Pharm Sci* 2018, **107**:1330–1341.
21. Hens B, Pathak SM, Mitra A, Patel N, Liu B, Patel S, Jamei M, Brouwers J, Augustijns P, Turner DB: ***In Silico* Modeling Approach for the Evaluation of Gastrointestinal Dissolution, Supersaturation, and Precipitation of Posaconazole**. *Mol Pharm* 2017, **14**:4321–4333.
22. Thiel C, Schneckener S, Krauss M, Ghallab A, Hofmann U, Kanacher T, Zellmer S, Gebhardt R, Hengstler JG, Kuepfer L: **A Systematic Evaluation of the Use of Physiologically Based Pharmacokinetic Modeling for Cross-Species Extrapolation**. *J Pharm Sci* 2015, **104**:191–206.
23. Chen Y, Zhao K, Liu F, Xie Q, Zhong Z, Miao M, Liu X, Liu L: **Prediction of Deoxypodophyllotoxin Disposition in Mouse, Rat, Monkey, and Dog by Physiologically Based Pharmacokinetic Model and the Extrapolation to Human**. *Front Pharmacol* 2016, **7**:488.
24. Campbell JL, Andersen ME, Hinderliter PM, Yi KD, Pastoor TP, Breckenridge CB, Clewell HJ: **PBPK Model for Atrazine and Its Chlorotriazine Metabolites in Rat and Human**. *Toxicol Sci* 2016, **150**:441–453.
25. Maldonado EM, Leoncikas V, Fisher CP, Moore JB, Plant NJ, Kierzek AM: **Integration of Genome Scale Metabolic Networks and Gene Regulation of Metabolic Enzymes With Physiologically Based Pharmacokinetics.** *CPT pharmacometrics Syst Pharmacol* 2017, **6**:732–746.
26. Thiel C, Cordes H, Conde I, Castell JV, Blank LM, Kuepfer L: **Model-based contextualization of in vitro toxicity data quantitatively predicts in vivo drug response in patients**. *Arch Toxicol* 2017, **91**:865–883.
27. Thiel C, Cordes H, Fabbri L, Aschmann HE, Baier V, Smit I, Atkinson F, Blank LM,

Kuepfer L: **A Comparative Analysis of Drug-Induced Hepatotoxicity in Clinically Relevant Situations**. *PLOS Comput Biol* 2017, **13**:e1005280.

28. Andersen ME, Black MB, Campbell JL, Pendse SN, Clewell III HJ, Pottenger LH, Bus JS, Dodd DE, Kemp DC, McMullen PD: **Combining transcriptomics and PBPK modeling indicates a primary role of hypoxia and altered circadian signaling in dichloromethane carcinogenicity in mouse lung and liver**. *Toxicol Appl Pharmacol* 2017, **332**:149–158.
29. Sharma RP, Schuhmacher M, Kumar V: **Developing integrated PBPK/PD coupled mechanistic pathway model (miRNA-BDNF): An approach towards system toxicology**. *Toxicol Lett* 2017, **280**:79–91.
30. Mason CL, Leedale J, Tasoulis S, Jarman I, Antoine DJ, Webb SD: **Systems Toxicology Approach to Identifying Paracetamol Overdose**. *CPT Pharmacometrics Syst Pharmacol* 2018, doi:10.1002/psp4.12298.
31. Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet J-P, Subramanian A, Ross KN, et al.: **The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease.** *Science* 2006, **313**:1929–35.
32. De Abrew KN, Kainkaryam RM, Shan YK, Overmann GJ, Settivari RS, Wang X, Xu J, Adams RL, Tiesman JP, Carney EW, et al.: **Grouping 34 Chemicals Based on Mode of Action Using Connectivity Mapping**. *Toxicol Sci* 2016, **151**:447–461.
33. Chen B, Ma L, Paik H, Sirota M, Wei W, Chua M-S, So S, Butte AJ: **Reversal of cancer gene expression correlates with drug efficacy and reveals therapeutic targets**. *Nat Commun* 2017, **8**:16022.
34. Liu X, Zeng P, Cui Q, Zhou Y: **Comparative analysis of genes frequently regulated by drugs based on connectivity map transcriptome data**. *PLoS One* 2017, **12**:e0179037.
35. El-Hachem N, Gendoo DMA, Ghoraie LS, Safikhani Z, Smirnov P, Chung C, Deng K, Fang A, Birkwood E, Ho C, et al.: **Integrative Cancer Pharmacogenomics to Infer Large-Scale Drug Taxonomy.** *Cancer Res* 2017, **77**:3057–3069.
36. Rempel E, Hoelting L, Waldmann T, Balmer N V., Schildknecht S, Grinberg M, Das Gaspar JA, Shinde V, Stöber R, Marchan R, et al.: **A transcriptome-based classifier to identify developmental toxicants by stem cell testing: design, validation and optimization for histone deacetylase inhibitors**. *Arch Toxicol* 2015, **89**:1599–1618.

37. Giordano M, Tripathi KP, Guarracino MR: **Ensemble of rankers for efficient gene signature extraction in smoke exposure classification**. *BMC Bioinformatics* 2018, **19**:48.
38. Wang Z, Monteiro CD, Jagodnik KM, Fernandez NF, Gundersen GW, Rouillard AD, Jenkins SL, Feldmann AS, Hu KS, McDermott MG, et al.: **Extraction and analysis of signatures from the Gene Expression Omnibus by the crowd**. *Nat Commun* 2016, **7**:12846.
39. Darde TA, Gaudriault P, Beranger R, Lancien C, Caillarec-Joly A, Sallou O, Bonvallot N, Chevrier C, Mazaud-Guittot S, Jégou B, et al.: **TOXsIgN: a cross-species repository for toxicogenomic signatures**. *Bioinformatics* 2018, doi:10.1093/bioinformatics/bty040.
40. Fabregat A, Jupe S, Matthews L, Sidiropoulos K, Gillespie M, Garapati P, Haw R, Jassal B, Korninger F, May B, et al.: **The Reactome Pathway Knowledgebase**. *Nucleic Acids Res* 2018, **46**:D649–D655.
41. Slenter DN, Kutmon M, Hanspers K, Riutta A, Windsor J, Nunes N, Mélius J, Cirillo E, Coort SL, Digles D, et al.: **WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research**. *Nucleic Acids Res* 2018, **46**:D661–D667.
42. Consortium TGO: **Expansion of the Gene Ontology knowledgebase and resources**. *Nucleic Acids Res* 2017, **45**:D331–D338.
43. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP: **Molecular signatures database (MSigDB) 3.0.** *Bioinformatics* 2011, **27**:1739–1740.
44. Parmentier C, Couttet P, Wolf A, Zaccharias T, Heyd B, Bachellier P, Uteng M, Richert L: **Evaluation of transcriptomic signature as a valuable tool to study drug-induced cholestasis in primary human hepatocytes**. *Arch Toxicol* 2017, **91**:2879–2893.
45. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al.: **Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.** *Proc Natl Acad Sci U S A* 2005, **102**:15545–50.
46. Yang L, Allen BC, Thomas RS: **BMDExpress: a software tool for the benchmark dose analyses of genomic data.** *BMC Genomics* 2007, **8**:387.
47. Dean JL, Jay Zhao Q, Lambert JC, Hawkins BS, Thomas RS, Wesselkamper SC: **Application of Gene Set Enrichment Analysis for Identification of Chemically-**

Induced, Biologically Relevant Transcriptomic Networks and Potential Utilization in Human Health Risk Assessment. *Toxicol Sci* 2017, **157**:kfx021.

48. Sirci F, Napolitano F, Pisonero-Vaquero S, Carrella D, Medina DL, di Bernardo D: **Comparing structural and transcriptional drug networks reveals signatures of drug activity and toxicity in transcriptional responses**. *npj Syst Biol Appl* 2017, **3**:23.
49. Kohonen P, Parkkinen JA, Willighagen EL, Ceder R, Wennerberg K, Kaski S, Grafström RC: **A transcriptomics data-driven gene space accurately predicts liver cytopathology and drug-induced liver injury.** *Nat Commun* 2017, **8**:15932.
50. Hendrickx DM, Souza T, Jennen DGJ, Kleinjans JCS: **DTNI: a novel toxicogenomics data analysis tool for identifying the molecular mechanisms underlying the adverse effects of toxic compounds**. *Arch Toxicol* 2017, **91**:2343–2352.
51. Souza TM, Kleinjans JCS, Jennen DGJ: **Dose and Time Dependencies in Stress Pathway Responses during Chemical Exposure: Novel Insights from Gene Regulatory Networks**. *Front Genet* 2017, **8**:142.
52. Taškova K, Fontaine J-F, Mrowka R, Andrade-Navarro MA: **Evaluation of in vivo and in vitro models of toxicity by comparison of toxicogenomics data with the literature**. *Methods* 2018, **132**:57–65.
53. Csardi G, Kutalik Z, Bergmann S: **Modular analysis of gene expression data with R**. *Bioinformatics* 2010, **26**:1376–1377.
54. AbdulHameed MDM, Ippolito DL, Stallings JD, Wallqvist A: **Mining kidney toxicogenomic data by using gene co-expression modules**. *BMC Genomics* 2016, **17**:790.
55. Te JA, AbdulHameed MDM, Wallqvist A: **Systems toxicology of chemically induced liver and kidney injuries: histopathology-associated gene co-expression modules**. *J Appl Toxicol* 2016, **36**:1137–1149.
56. Pang W, Lian F-Z, Leng X, Wang S, Li Y, Wang Z, Li K, Gao Z, Jiang Y: **Microarray expression profiling and co-expression network analysis of circulating LncRNAs and mRNAs associated with neurotoxicity induced by BPA**. *Environ Sci Pollut Res* 2018, doi:10.1007/s11356-018-1678-y.
57. Zhang B, Horvath S: **A General Framework for Weighted Gene Co-Expression Network Analysis**. *Stat Appl Genet Mol Biol* 2005, **4**:Article17.

58. Sutherland JJ, Webster YW, Willy JA, Searfoss GH, Goldstein KM, Irizarry AR, Hall DG, Stevens JL: **Toxicogenomic module associations with pathogenesis: a network-based approach to understanding drug toxicity**. *Pharmacogenomics J* 2017, doi:10.1038/tpj.2017.17.

59. Ivliev AE, 't Hoen PAC, Borisevich D, Nikolsky Y, Sergeeva MG: **Drug Repositioning through Systematic Mining of Gene Coexpression Networks in Cancer**. *PLoS One* 2016, **11**:e0165059.

60. Vahle JL, Anderson U, Blomme EAG, Hoflack J-C, Stiehl DP: **Use of toxicogenomics in drug safety evaluation: Current status and an industry perspective**. *Regul Toxicol Pharmacol* 2018, doi:10.1016/J.YRTPH.2018.04.011.

61. Fielden MR, Zacharewski TR: **Challenges and Limitations of Gene Expression Profiling in Mechanistic and Predictive Toxicology**. *Toxicol Sci* 2001, **60**:6–10.

62. Berger SI, Iyengar R: **Role of systems pharmacology in understanding drug adverse events.** *Wiley Interdiscip Rev Syst Biol Med* 2011, **3**:129–35.

63. NOBLE D: **Cardiac Action and Pacemaker Potentials based on the Hodgkin-Huxley Equations**. *Nature* 1960, **188**:495–497.

64. O'Hara T, Virág L, Varró A, Rudy Y: **Simulation of the Undiseased Human Cardiac Ventricular Action Potential: Model Formulation and Experimental Validation**. *PLoS Comput Biol* 2011, **7**:e1002061.

65. Colatsky T, Fermini B, Gintant G, Pierson JB, Sager P, Sekino Y, Strauss DG, Stockbridge N: **The Comprehensive in Vitro Proarrhythmia Assay (CiPA) initiative — Update on progress**. *J Pharmacol Toxicol Methods* 2016, **81**:15–20.

66. Yang P-C, Perissinotti LL, López-Redondo F, Wang Y, DeMarco KR, Jeng M-T, Vorobyov I, Harvey RD, Kurokawa J, Noskov SY, et al.: **A multiscale computational modelling approach predicts mechanisms of female sex risk in the setting of arousal-induced arrhythmias**. *J Physiol* 2017, **595**:4695–4723.

67. Li Z, Dutta S, Sheng J, Tran PN, Wu W, Chang K, Mdluli T, Strauss DG, Colatsky T: **Improving the In Silico Assessment of Proarrhythmia Risk by Combining hERG (Human Ether-à-go-go-Related Gene) Channel-Drug Binding Kinetics and Multichannel Pharmacology.** *Circ Arrhythm Electrophysiol* 2017, **10**:e004628.

68. Passini E, Britton OJ, Lu HR, Rohrbacher J, Hermans AN, Gallacher DJ, Greig RJH,

Bueno-Orovio A, Rodriguez B: **Human In Silico Drug Trials Demonstrate Higher Accuracy than Animal Models in Predicting Clinical Pro-Arrhythmic Cardiotoxicity**. *Front Physiol* 2017, **8**:668.

69. Romero L, Cano J, Gomis-Tena J, Trenor B, Sanz F, Pastor M, Saiz J: **In Silico QT and APD Prolongation Assay for Early Screening of Drug-Induced Proarrhythmic Risk**. *J Chem Inf Model* 2018, **58**:867–878.
70. Bloomingdale P, Nguyen VA, Niu J, Mager DE: **Boolean network modeling in systems pharmacology**. *J Pharmacokinet Pharmacodyn* 2018, **45**:159–180.
71. Traynard P, Tobalina L, Eduati F, Calzone L, Saez-Rodriguez J: **Logic Modeling in Quantitative Systems Pharmacology.** *CPT pharmacometrics Syst Pharmacol* 2017, **6**:499–511.
72. Le Novère N: **Quantitative and logic modelling of molecular and gene networks**. *Nat Rev Genet* 2015, **16**:146–158.
73. Gómez Tejeda Zañudo J, Scaltriti M, Albert R: **A network modeling approach to elucidate drug resistance mechanisms and predict combinatorial drug treatments in breast cancer**. *Cancer Converg* 2017, **1**:5.
74. Zhu P, Aliabadi HM, Uludağ H, Han J: **Identification of Potential Drug Targets in Cancer Signaling Pathways using Stochastic Logical Models**. *Sci Rep* 2016, **6**:23078.
75. Poret A, Guziolowski C: **Therapeutic target discovery using Boolean network attractors: improvements of kali**. *R Soc Open Sci* 2018, **5**:171852.
76. Vitali F, Cohen LD, Demartini A, Amato A, Eterno V, Zambelli A, Bellazzi R: **A Network-Based Data Integration Approach to Support Drug Repurposing and Multi-Target Therapies in Triple Negative Breast Cancer**. *PLoS One* 2016, **11**:e0162407.
77. Morris M, Clarke D, Osimiri L, Lauffenburger D: **Systematic Analysis of Quantitative Logic Model Ensembles Predicts Drug Combination Effects on Cell Signaling Networks**. *CPT Pharmacometrics Syst Pharmacol* 2016, **5**:544–553.
78. Samaga R, Klamt S: **Modeling approaches for qualitative and semi-quantitative analysis of cellular signaling networks**. *Cell Commun Signal* 2013, **11**:43.
79. Melas IN, Sakellaropoulos T, Iorio F, Alexopoulos LG, Loh W-Y, Lauffenburger DA, Saez-Rodriguez J, Bai JPF: **Identification of drug-specific pathways based on gene**

expression data: application to drug induced lung injury. *Integr Biol (Camb)* 2015, **7**:904–20.

80. Messinis DE, Melas IN, Hur J, Varshney N, Alexopoulos LG, Bai JPF: **Translational Systems Pharmacology-Based Predictive Assessment of Drug-Induced Cardiomyopathy**. *CPT Pharmacometrics Syst Pharmacol* 2018, **7**:166–174.
81. Gille C, Bölling C, Hoppe A, Bulik S, Hoffmann S, Hübner K, Karlstädt A, Ganeshan R, König M, Rother K, et al.: **HepatoNet1: a comprehensive metabolic reconstruction of the human hepatocyte for the analysis of liver physiology.** *Mol Syst Biol* 2010, **6**:411.
82. Thiele I, Swainston N, Fleming RMT, Hoppe A, Sahoo S, Aurich MK, Haraldsdottir H, Mo ML, Rolfsson O, Stobbe MD, et al.: **A community-driven global reconstruction of human metabolism**. *Nat Biotechnol* 2013, **31**:419–425.
83. Chang RL, Xie L, Xie L, Bourne PE, Palsson BØ: **Drug Off-Target Effects Predicted Using Structural Analysis in the Context of a Metabolic Network Model**. *PLoS Comput Biol* 2010, **6**:e1000938.
84. Mardinoglu A, Agren R, Kampf C, Asplund A, Uhlen M, Nielsen J: **Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease**. *Nat Commun* 2014, **5**:3083.
85. Mardinoglu A, Agren R, Kampf C, Asplund A, Nookaew I, Jacobson P, Walley AJ, Froguel P, Carlsson LM, Uhlen M, et al.: **Integration of clinical data with a genome-scale metabolic model of the human adipocyte.** *Mol Syst Biol* 2013, **9**:649.
86. Sahoo S, Thiele I: **Predicting the impact of diet and enzymopathies on human small intestinal epithelial cells**. *Hum Mol Genet* 2013, **22**:2705–2722.
87. Väremo L, Scheele C, Broholm C, Mardinoglu A, Kampf C, Asplund A, Nookaew I, Uhlén M, Pedersen BK, Nielsen J: **Proteome- and Transcriptome-Driven Reconstruction of the Human Myocyte Metabolic Network and Its Use for Identification of Markers for Diabetes**. *Cell Rep* 2015, **11**:921–933.
88. Ravikrishnan A, Raman K: **Critical assessment of genome-scale metabolic networks: the need for a unified standard**. *Brief Bioinform* 2015, **16**:1057–1068.
89. Blais EM, Rawls KD, Dougherty B V, Li ZI, Kolling GL, Ye P, Wallqvist A, Papin JA: **Reconciled rat and human metabolic networks for comparative toxicogenomics and biomarker predictions.** *Nat Commun* 2017, **8**:14250.

90. Carbonell P, Lopez O, Amberg A, Pastor M, Sanz F: **Hepatotoxicity prediction by systems biology modeling of disturbed metabolic pathways using gene expression data**. *ALTEX* 2016, **34**:219–234.
91. Sier JH, Thumser AE, Plant NJ: **Linking physiologically-based pharmacokinetic and genome-scale metabolic networks to understand estradiol biology**. *BMC Syst Biol* 2017, **11**:141.
92. Shaked I, Oberhardt MA, Atias N, Sharan R, Ruppin E: **Metabolic Network Prediction of Drug Side Effects**. *Cell Syst* 2016, **2**:209–213.
93. Hoekstra A, Chopard B, Coveney P: **Multiscale modelling and simulation: a position paper.** *Philos Trans A Math Phys Eng Sci* 2014, **372**:20130377.
94. Vicini P: **Multiscale Modeling in Drug Discovery and Development: Future Opportunities and Present Challenges**. *Clin Pharmacol Ther* 2010, **88**:126–129.
95. Sluka JP, Fu X, Swat M, Belmonte JM, Cosmanescu A, Clendenon SG, Wambaugh JF, Glazier JA: **A Liver-Centric Multiscale Modeling Framework for Xenobiotics**. *PLoS One* 2016, **11**:e0162428.
96. Cordes H, Thiel C, Baier V, Blank LM, Kuepfer L: **Integration of genome-scale metabolic networks into whole-body PBPK models shows phenotype-specific cases of drug-induced metabolic perturbation**. *npj Syst Biol Appl* 2018, **4**:10.
97. Sahli Costabal F, Yao J, Kuhl E: **Predicting drug-induced arrhythmias by multiscale modeling**. *Int j numer method biomed eng* 2018, **34**:e2964.
98. Ribba B, Grimm HP, Agoram B, Davies MR, Gadkar K, Niederer S, van Riel N, Timmis J, van der Graaf PH: **Methodologies for Quantitative Systems Pharmacology (QSP) Models: Design and Estimation.** *CPT pharmacometrics Syst Pharmacol* 2017, **6**:496–498.
99. Breschi A, Gingeras TR, Guigó R: **Comparative transcriptomics in human and mouse**. *Nat Rev Genet* 2017, **18**:425–440.
100. El-Hachem N, Grossmann P, Blanchet-Cohen A, Bateman AR, Bouchard N, Archambault J, Aerts HJWL, Haibe-Kains B: **Characterization of Conserved Toxicogenomic Responses in Chemically Exposed Hepatocytes across Species and Platforms.** *Environ Health Perspect* 2016, **124**:313–20.
101. Sutherland JJ, Jolly RA, Goldstein KM, Stevens JL: **Assessing Concordance of Drug-**

Induced Transcriptional Response in Rodent Liver and Cultured Hepatocytes. *PLOS Comput Biol* 2016, **12**:e1004847.

102. El-Hachem N, Grossmann P, Blanchet-Cohen A, Bateman AR, Bouchard N, Archambault J, Aerts HJWL, Haibe-Kains B: **Characterization of Conserved Toxicogenomic Responses in Chemically Exposed Hepatocytes across Species and Platforms.** *Environ Health Perspect* 2016, **124**:313–20.
103. Shankaran H, Cronin A, Barnes J, Sharma P, Tolsma J, Jasper P, Mettetal JT: **Systems Pharmacology Model of Gastrointestinal Damage Predicts Species Differences and Optimizes Clinical Dosing Schedules**. *CPT Pharmacometrics Syst Pharmacol* 2018, **7**:26–33.
104. Mih N, Brunk E, Bordbar A, Palsson BO: **A Multi-scale Computational Platform to Mechanistically Assess the Effect of Genetic Variation on Drug Responses in Human Erythrocyte Metabolism.** *PLoS Comput Biol* 2016, **12**:e1005039.
105. Lek M, Karczewski KJ, Minikel E V., Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, et al.: **Analysis of protein-coding genetic variation in 60,706 humans**. *Nature* 2016, **536**:285–291.
106. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, Chakravarti A, Clark AG, Donnelly P, Eichler EE, et al.: **A global reference for human genetic variation**. *Nature* 2015, **526**:68–74.
107. Turnbull C, Scott RH, Thomas E, Jones L, Murugaesu N, Pretty FB, Halai D, Baple E, Craig C, Hamblin A, et al.: **The 100 000 Genomes Project: bringing whole genome sequencing to the NHS.** *BMJ* 2018, **361**:k1687.
108. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, Burgess S, Jiang T, Paige E, Surendran P, et al.: **Genomic atlas of the human plasma proteome**. *Nature* 2018, **558**:73–79.
109. Sanz F, Pognan F, Steger-Hartmann T, Díaz C, eTOX, Cases M, Pastor M, Marc P, Wichard J, Briggs K, et al.: **Legacy data sharing to improve drug safety assessment: the eTOX project**. *Nat Rev Drug Discov 2017 1612* 2017,
110. Fisel P, Schaeffeler E, Schwab M: **DNA Methylation of ADME Genes**. *Clin Pharmacol Ther* 2016, **99**:512–527.
111. Koppel N, Maini Rekdal V, Balskus EP: **Chemical transformation of xenobiotics by the**

human gut microbiota. *Science* 2017, **356**:eaag2770.

112. Wilson ID, Nicholson JK: **Gut microbiome interactions with drug metabolism, efficacy, and toxicity**. *Transl Res* 2017, **179**:204–222.
113. Magnúsdóttir S, Heinken A, Kutt L, Ravcheev DA, Bauer E, Noronha A, Greenhalgh K, Jäger C, Baginska J, Wilmes P, et al.: **Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota**. *Nat Biotechnol* 2016, **35**:81–89.
114. Thiele I, Clancy CM, Heinken A, Fleming RMT: **Quantitative systems pharmacology and the personalized drug–microbiota–diet axis**. *Curr Opin Syst Biol* 2017, **4**:43–52.
115. OECD: *Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models*. 2014.
116. Igarashi Y, Nakatsu N, Yamashita T, Ono A, Ohno Y, Urushidani T, Yamada H: **Open TG-GATEs: a large-scale toxicogenomics database**. *Nucleic Acids Res* 2015, **43**:D921–D927.
117. Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, Gould J, Davis JF, Tubelli AA, Asiedu JK, et al.: **A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles.** *Cell* 2017, **171**:1437–1452.e17.
118. Lamb J: **The Connectivity Map: a new tool for biomedical research.** *Nat Rev Cancer* 2007, **7**:54–60.
119. Shoemaker RH: **The NCI60 human tumour cell line anticancer drug screen**. *Nat Rev Cancer* 2006, **6**:813–823.
120. Keenan AB, Jenkins SL, Jagodnik KM, Koplev S, He E, Torre D, Wang Z, Dohlman AB, Silverstein MC, Lachmann A, et al.: **The Library of Integrated Network-Based Cellular Signatures NIH Program: System-Level Cataloging of Human Cells Response to Perturbations**. *Cell Syst* 2018, **6**:13–24.
121. Lea IA, Gong H, Paleja A, Rashid A, Fostel J: **CEBS: a comprehensive annotated database of toxicological data**. *Nucleic Acids Res* 2017, **45**:D964–D971.

Annotated Papers

25.** Maldonado EM, Leoncikas V, Fisher CP, Moore JB, Plant NJ, Kierzek AM: **Integration of Genome Scale Metabolic Networks and Gene Regulation of Metabolic Enzymes With Physiologically Based Pharmacokinetics.** *CPT pharmacometrics Syst Pharmacol* 2017, **6**:732–746

This review describes the current status of multiscale models combining PBPK, GSMN, and gene regulatory networks, and provides an example on how to build a multiscale model integrating HepatoNet1, (a liver GSMN), the gene regulatory network of cytochrome CYP3A4, and a whole body PBPK model, and two example applications.

26.* Thiel C, Cordes H, Conde I, Castell JV, Blank LM, Kuepfer L: **Model-based contextualization of in vitro toxicity data quantitatively predicts in vivo drug response in patients**. *Arch Toxicol* 2017, **91**:865–883.

The authors develop a multiscale aproach that combines PBPK models with toxicogenomics data that is able to translate *in vitro* drug response data to the *in vivo* context. As example they use the approach to predict *in vivo* rat and patient responses to acute azathioprine overdose.

30.* Mason CL, Leedale J, Tasoulis S, Jarman I, Antoine DJ, Webb SD: **Systems Toxicology Approach to Identifying Paracetamol Overdose**. *CPT Pharmacometrics Syst Pharmacol* 2018, doi:10.1002/psp4.12298.

The authors propose an *in silico* paracetamol toxicity identification framework that uses a PBPK model of paracetamol and takes into account a panel of biomarkers of liver injury. The framework allows to predict the paracetamol dose, and time since administration, as well as the probability of DILI in mice models.

35.* El-Hachem N, Gendoo DMA, Ghoraie LS, Safikhani Z, Smirnov P, Chung C, Deng K, Fang A, Birkwood E, Ho C, et al.: **Integrative Cancer Pharmacogenomics to Infer Large-Scale Drug Taxonomy.** *Cancer Res* 2017, **77**:3057–3069.

The authors develop a computational method that combines different types of drug-centric data: structural information, toxicogenomics data, and growth inhibition profiles from cancer cell lines. From each type of data, drug–drug similarity matrices are built, which are combined using a network fusion algorithm to produce a drug taxonomy that succesfully classifies drugs in major classes, and predicts new drug targets

49.* Kohonen P, Parkkinen JA, Willighagen EL, Ceder R, Wennerberg K, Kaski S, Grafström RC: **A transcriptomics data-driven gene space accurately predicts liver cytopathology and drug-induced liver injury.** *Nat Commun* 2017, **8**:15932.

The authors apply an approach to reduce dimentionality of toxicogenomics data, and combine it with cytotoxicity data to produce the PTGS, a collection of 14 gene sets that can be used to predict liver toxicity, and group drugs by mode of action, among other examples.

50.* Hendrickx DM, Souza T, Jennen DGJ, Kleinjans JCS: **DTNI: a novel toxicogenomics data analysis tool for identifying the molecular mechanisms underlying the adverse effects of toxic compounds**. *Arch Toxicol* 2017, **91**:2343–2352.

The authors develop an ODE-based method that infers gene regulatory networks using time and dose toxicogenomics data, and a set of genes belonging to a process of interest. The method is able to provide statistic significance for the interactions among the genes in the gene set, and to pinpoint compound affecting the process under study.

58.** Sutherland JJ, Webster YW, Willy JA, Searfoss GH, Goldstein KM, Irizarry AR, Hall DG, Stevens JL: **Toxicogenomic module associations with pathogenesis: a network-based approach to understanding drug toxicity**. *Pharmacogenomics J* 2017, doi:10.1038/tpj.2017.17.

Using WGCNA to analyze gene expression data from rat liver, the authors produce the TGX map, a collection of modules of co-expressed genes, and show their utility to understand the mechanisms underlying specific liver toxicity phenotypes using several examples.

68.** Passini E, Britton OJ, Lu HR, Rohrbacher J, Hermans AN, Gallacher DJ, Greig RJH, Bueno-Orovio A, Rodriguez B: **Human In Silico Drug Trials Demonstrate Higher Accuracy than Animal Models in Predicting Clinical Pro-Arrhythmic Cardiotoxicity**. *Front Physiol* 2017, **8**:668.

The authors show that *in silico* cardiac electrophysiological models exhibit more accurate risk prediction of clinical pro-arrhythmic cardiotoxicity compared to animal experiments and human induced pluripotent stem cell-derived cardiomyocytes.

69.** Romero L, Cano J, Gomis-Tena J, Trenor B, Sanz F, Pastor M, Saiz J: **In Silico QT and APD Prolongation Assay for Early Screening of Drug-Induced Proarrhythmic Risk**. *J Chem Inf Model* 2018, **58**:867–878.

The authors show that *in silico* simulations that combine drug concentration, and multiple

channel–drug interactions in isolated endocardial, midmyocardial, and epicardial cells outperform the standard potassium hERG channels block test in the taks of predictiong drug-induced proarrhythmic risk.

79.* Melas IN, Sakellaropoulos T, Iorio F, Alexopoulos LG, Loh W-Y, Lauffenburger DA, Saez-Rodriguez J, Bai JPF: **Identification of drug-specific pathways based on gene expression data: application to drug induced lung injury.** *Integr Biol (Camb)* 2015, **7**:904–20.

The authors develop a framework that combines toxicogenomics data, drug targets and transcription factor information, and a global signaling network to find specific pathways underlying the mode of action of a drug. As a proof of concept, they use drugs known to produce lung toxicity, and produce a signaling subnetwork capable of explaining the mechanisms underlying drug-induced lung injury

103* Shankaran H, Cronin A, Barnes J, Sharma P, Tolsma J, Jasper P, Mettetal JT: **Systems Pharmacology Model of Gastrointestinal Damage Predicts Species Differences and Optimizes Clinical Dosing Schedules**. *CPT Pharmacometrics Syst Pharmacol* 2018, **7**:26–33.

The authors build a matemathical model that describe the dynamics of different types of intestinal cells to predict gastrointestinal toxicity of irinotecan. The model was fit using measurements in rats, and was able to translate into quantitative predictions of enterocyte loss and recovery kinetics in humans.

104** Mih N, Brunk E, Bordbar A, Palsson BO: **A Multi-scale Computational Platform to Mechanistically Assess the Effect of Genetic Variation on Drug Responses in Human Erythrocyte Metabolism.** *PLoS Comput Biol* 2016, **12**:e1005039.

An example of how protein structural information can be integrated to GSMN models to predict the impact of genetic variation on drug metabolism.

109** Sanz F, Pognan F, Steger-Hartmann T, Díaz C, eTOX, Cases M, Pastor M, Marc P, Wichard J, Briggs K, et al.: **Legacy data sharing to improve drug safety assessment: the eTOX project.** *Nat Rev Drug Discov* 2017, 1612 2017

The authors discuss their experience in building a collaborative model for sharing legacy pre-clinical safety data among pharmaceutical companies, leading to the development of the eTOX database and facilitating the development of predictive models.