ELSEVIER

The value of *in silico* chemistry in the safety assessment of chemicals in the consumer goods and pharmaceutical industries

Sandeep Modi, Michael Hughes, Andrew Garrow and Andrew White

Safety & Environmental Assurance Centre, Unilever R&D, Colworth Science Park, Sharnbrook, Bedford, England, UK MK44 1LQ

In silico **toxicology prediction is an extremely challenging area because many toxicological effects are a result of changes in multiple physiological processes. In this article we discuss limitations and strengths of these *in silico* tools. Additionally, we look at different parameters that are necessary to make the best use of these tools, and also how to gain acceptance outside the modelling community and into the regulatory arena. As a solution, we propose an integrated workflow for combined use of data extraction, quantitative structure activity relationships and read-across methods. We also discuss how the recent advances in this field can enable transition to a new paradigm of the discovery process, as exemplified by the Toxicity Testing in the 21st Century initiative.**

Introduction

Chemicals in drugs, food and consumer goods are an integral part of our everyday life whether they be 'naturals' or manmade. Although the regulatory needs and requirements for risk assessments differ between consumer goods and pharmaceutical industries, it is evident that the *in silico* approaches currently available can offer significant benefit to both of these sectors.

The well recognised definition of a good and successful drug is an appropriate balance of potency, efficacy, safety and favourable pharmacokinetics. For the drug industry, it is imperative that a chemical can reach its target with a concentration at the target site suitable for the desired efficacy and also below toxicity thresholds. To avoid late-stage failures in the discovery of new chemicals to be used as drugs, absorption, distribution, metabolism, elimination and toxicity (ADMET) studies are now mostly carried out at a much earlier stage of the discovery process [1–3]. However, owing to the use of combinatorial chemistry libraries and high-throughput screening (HTS), there has been a dramatic increase in the number of active chemicals. Consequently, the discovery process has struggled to keep up with this increase in influx of chemicals. Similarly within the consumer goods industry, late-stage failures are costly in terms of resource and time due to safety concerns and as such there is a strong rationale to bring safety assessments closer to the initial stages of the innovation pipeline where a larger pool of candidate compounds is available. Usually, the traditional safety methods centred around toxicological assessment and testing are unable to cope with the influx of new chemicals to be tested in a timely way. There is also growing interest in the development of alternative approaches to toxicity testing [4] that reduce, refine or replace the use of animals in safety assessment and also enable faster high throughput assessment of hazard and risk that is relevant to human safety assessment. To address this, there is a need for development of both *in silico* virtual models and also a better understanding of the effects of different chemicals on physiological processes. This could then be deployed for toxicological assessments at an earlier stage of the discovery process and for assessment of larger numbers of chemicals [5–9].

Why *in silico*?

In silico screening is typically a low cost high-throughput process, which can provide a fast indication of potential hazards for use in lead prioritisation [6–9]. As no physical compounds are required, these screens can be run on virtual compounds at early stages of discovery to prioritise chemicals for ADMET testing. Additionally, *in silico* tools can help provide a mechanistic understanding of these predictions, for example, to explain why a compound is predicted to be active or inactive. This information can then be used to re-engineer a chemical, alter its ADMET profile or design out the toxicity of new chemicals. These predictive models can be built either directly on data from *in vitro* assays (e.g. Ames bacterial

Corresponding author:. Modi, S. (sandeep.modi@unilever.com)

1359-6446/06/$ - see front matter doi:10.1016/j.drudis.2011.10.022

Please cite this article in press as: S.. Modi, et al., The value of *in silico* chemistry in the safety assessment of chemicals in the consumer goods and pharmaceutical industries, Drug Discov Today (2011), doi:10.1016/j.drudis.2011.10.022

mutation assays) or directly on *in vivo* data (e.g. carcinogenicity, TD50s), and can also be used for understanding of *in vitro* to *in vivo* extrapolation. Following successful validation of a predictive model, it might be possible to have faster cycle times, lower costs, and early indication of drug failure, with a much reduced need for *in vitro* or animal testing. Depending upon the toxicology endpoint, some of these *in silico* tools and technologies are considered to be valid and are recognised by some regulatory agencies (e.g. for use in the European REACH initiative) [10]. A set of principles from the Organisation for Economic Co-operation and Development (OECD) must be followed to achieve regulatory acceptance of the predictive chemistry tools. These include questions around the defined endpoint (i.e. data used for modelling, applicability domain, methods, their mechanistic interpretation and appropriate measures of predictivity [11]). In the remainder of this article, we discuss the importance of these points, and why it is crucial to get them right to gain the full benefits of *in silico* chemistry.

Choosing the right data

Because toxic effects are still responsible for some 20% of the late-stage failures in drug development, there is an urgent need for *in silico* tools that can be used to estimate the toxicology profile of a chemical [6]. Typically, adverse reactions are not discovered until after market release. This is an expensive scenario that is all too common in the discovery of new drugs and chemicals. The most relevant data usually comes from human clinical trials, which can be very expensive to generate and also limited in number of data points. The data chosen for modelling toxicity has to be chosen wisely for any given toxicology endpoint. For example, in the case of genetic toxicity, there is a large body of *in vitro* data available from Ames testing compared with limited *in vivo* data (e.g. carcinogenicity in different species). While the *in vitro* data provides a comprehensive model training set, it must be noted that the final goal of *in silico* models is to predict the *in vivo* effects in humans for a given chemical. One must therefore choose *in vitro* toxicology endpoints carefully keeping its relationship and relevance to the *in vivo* data in mind. It is clear that the use of animals has limitations, for example, humans are not 70 kg rats, we absorb/metabolise chemicals differently; we live longer (enabling certain diseases to develop, prompting evolutionary adaptations to protect against them); and we are exposed to a multitude of environmental factors [12,13]. The models are clearly only as good as the data they are based on, there is still no replacement for the expression 'garbage in, garbage out' [14]. Before modelling any data, the modeller needs to ensure the quality of the datasets from different sources and standardisation procedures must be in place to cope with data in different formats. With the recent advances in HTS, chemical synthesis and biological screening, there is no shortage of publicly or commercially available databases that can be used as data sources for these models [15]. A recent article discusses some of these issues and has also evaluated the consequences of both random and systematic errors with chemical structure curations in well-known datasets [16]. Some useful electronic resources that contain data suitable for toxicology model building are listed in Table 1.

TABLE 1

List of public data sources as useful training sets for predictive chemistry models

Database	Brief description	Refs
ACToR – Aggregated Computational Toxicology Resource	ACToR (Aggregated Computational Toxicology Resource) is a collection of databases collated or developed by the US EPA National Center for Computational Toxicology (NCCT). Data includes chemical structure, physico-chemical values, *in vitro* assay data, expo	ACToR: http://actor.epa.gov/actor/actor_help_20080903.htm
CCRIS	Chemical Carcinogenesis Research Information System – carcinogenicity, mutagenicity, tumor promotion, and tumor inhibition data provided by the National Cancer Institute (NCI)	U.S. National Library of Medicine: http://www.nlm.nih.gov/pubs/factsheets/ccrisfs.html
ChEMBL	Contains calculated properties (e.g. log *P*, molecular weight, Lipinski parameters, among others) and abstracted bioactivities (e.g. binding constants, pharmacology and ADMET data)	ChEMBL: https://www.ebi.ac.uk/chembldb/
Comparative Toxicogenomics Database (CTD)	Find associations between gene/proteins, environmental chemicals and toxicology	Comparative Toxicogenomics Database: http://ctd.mdibl.org/
CPDB (The Carcinogenic Potency Database)	Provides a broad perspective on possible cancer hazards from human exposures to chemicals that cause cancer in high dose rodent cancer tests	The Carcinogenic Potency Project: http://potency.berkeley.edu/
DART	Developmental and Reproductive Toxicology and Environmental Teratology Information Center – current and older literature on developmental and reproductive toxicology	U.S. National Library of Medicine: http://www.nlm.nih.gov/pubs/factsheets/dartfs.html
NTP (National Toxicology Program)	It contains toxicity studies from shorter duration tests and from genetic toxicity studies, which includes both *in vitro* and *in vivo* tests. It also contains the immunotoxicity, developmental toxicity and reproductive toxicity studies	National Toxicology Program: http://ntp-apps.niehs.nih.gov/ntp_tox/
RepDOSE	Repeat dose study data for dog, mouse and rat. Shows effects of chemicals on target organs. Studies are rated by reliability	RepDOSE: http://www.fraunhofer-repdose.de/
ToxRefDB	ToxRefDB (Toxicity Reference Database) captures thousands of *in vivo* animal toxicity studies on hundreds of chemicals	EPA: http://www.epa.gov/ncct/toxrefdb/

Please cite this article in press as: S.. Modi, et al., The value of *in silico* chemistry in the safety assessment of chemicals in the consumer goods and pharmaceutical industries, Drug Discov Today (2011), doi:10.1016/j.drudis.2011.10.022

Choice of methods

The three main predictive chemistry approaches, such as structure activity relationships (SAR), quantitative structure activity relationships (QSAR) and read-across, have been used in the past for the prediction of several toxicology endpoints.

SAR

This approach is associated with the local reactivity of chemicals and includes reactivity of functional groups also called structural alerts, pharmacophores or toxicophores [17–20]. This technique can be an invaluable tool in the *in silico* prediction of toxicity because it is simple and easy to understand, and highlights the presence of certain substructures within the molecule that can be related to an adverse reaction. Also, this alert approach can help provide mechanistic understanding of an observed adverse outcome. SAR usually works well with binary classifications and has been used for highlighting fragments of several toxicology endpoints, including Ames, carcinogenicity, hepatotoxicity and skin sensitisation [20].

QSAR

This approach provides the statistical relationship between the toxicity of a chemical and its physicochemical properties and structural characteristics. Different QSAR and machine learning methods have different ways of deriving these approximations to provide information about the toxic effect of chemicals [21]. Like SAR, this approach can work with binary classification [22–24] but also works with continuous data, such as rodent carcinogenicity and chronic toxicity data [25,26].

Read-across

Read-across of hazard data is a well recognised method for predicting the hazard profile of a substance where endpoint data are lacking by linking it to structurally similar compounds for which experimental data are available for a given endpoint [27,28]. This enables for a read-across approach to be used to predict the toxicity of those members of a chemical family for which no direct toxicology data are available.

To fully capitalise on the opportunities presented by these *in silico* tools, they need to be transparent and provide as much support and confidence behind each prediction as possible. As all biological endpoints (especially toxicity) are often the combination of multiple phenomena, most *in silico* models end up being complex [21,25,26]. There has to be balance between model interpretation, complexity and its predictive power. Depending on the endpoint, probabilistic and statistical QSAR methods (involving complex algorithms) might provide superior results as compared with simple SAR methods as it has been shown in the case of genotoxicity modelling [22–24]. The modelling algorithm linking molecular descriptors to the output variable needs to be chosen so that it takes the complexity of the particular relationship into account, otherwise overfitting (in case a complex modelling procedure is used) or insufficient predictivity of a model (in case a simple modelling procedure is used) might result. In case a complex algorithm is used in building the models, users might think that the models are an 'algorithmic jungle' and consequently, the benefits of the models could be easily misunderstood [29]. The precise choice of descriptors and modelling methods for a given model is generally dictated by the complexity of the process that underlies the given toxicity event, regulatory considerations, such as the OECD QSAR guidelines and the personal preference of the model developer. For example, for the drug-induced phospholipidosis, the potentially toxic excessive accumulation of phospholipids in cells/tissues might be described with simple descriptors, such as the presence of a positive charge/basic substituent and high lipophilicity [30]. By contrast, for example P450 inhibition or hERG binding are complex receptor-mediated processes, which arguably require complex non-linear modelling methods [31–33]. However, like with any complex problem, the task of building an understanding is much easier if this can be broken into different simple processes. A study [34] shows and suggests how *in silico* models for hERG inhibition can be used as early screening tools for eliminating potent hERG inhibitors from chemical libraries in early drug discovery. This could serve as an alternative to the more expensive and time consuming experimental assessments, or the model could be used to prioritise sending predicted inhibitors for experimental assay [34,35]. The choice of method, its transparency and its mechanistic interpretation might have an important role in getting the full benefits of these approaches, and also the acceptance of these tools outside the modelling community and into the regulatory arena. Table 2 lists some of the popular free open source software for prediction of various toxicity endpoints. The methods used in this table varies from simple SAR methods to complex QSAR methods.

Importance of applicability domain

In silico models might not perform well if a predicted chemical is beyond the chemical space where the models were developed [36–38]. Therefore, applicability domain is one of the main reasons for the QSAR/SAR model failure owing to the difference in chemical space of compounds that were used to develop and apply the models. This leads to the issue of whether global or local models should be used. Global models usually contain a large and diverse set of chemicals. These are generally suited to dealing with non-congeneric structural data and also when semi-quantitative predictions are needed. An example is discussed in a recent study for Ames test predictions [24]. By contrast, local models are built on a particular chemical series containing a small set of closely related chemicals and hence these local models might have a small applicability domain. An example of this has been shown for the Ames predictions for the aromatic amines [39]. *In silico* modellers should therefore highlight the validity of these *in silico* models by selecting and obtaining experimental data from new compound sets with structures different from those in the original model training sets [40]. For a bad model, accuracy measures tend to be biased towards the training set, but its performance then decreases when tested with a new set of chemicals [19].

How to make best use of (Q)SAR methods

It is important that the models built are continuously validated and refined based on new data and understanding the cause of the apparent 'failure' cases. It is well known that each individual model has its own problems and pitfalls [41–43]. For example, in the case of the expert derived structural alert approach [17–20], where the presence of a small fragment can be correlated with a

Please cite this article in press as: S.. Modi, et al., The value of *in silico* chemistry in the safety assessment of chemicals in the consumer goods and pharmaceutical industries, Drug Discov Today (2011), doi:10.1016/j.drudis.2011.10.022

TABLE 2

List of free open source tools/softwares for predictive toxicity models

Name of software	Brief description	Refs
CAESAR	Models for mutagenicity/carcinogenicity were developed and released as an open source software tool in the frame of the EU CAESAR project. In this two complementary approaches (regression and classification) were applied to develop models for carcinogenic	Computer Assisted Evaluation of industrial chemical Substances According to Regulations: http://www.caesar-project.eu/
Lazar	For predictions of Ames mutagenicity, carcinogenicity	Lazar Toxicity Predictions: http://lazar.in-silico.de
OECD Toolbox	For prediction of several end points as well as experimental data that can be used to support grouping and read-across	Oasis: http://toolbox.oasis-lmc.org
OncoLogic™	For predictions of carcinogenicity	http://www.epa.gov/oppt/sf/pubs/oncologic.htm
Tox-Comp	Flexible, modular system for the early assessment of the cardiotoxic potency	Tox-Comp.net: http://tox-comp.net/
ToxTree	This was developed by Ideaconsult Ltd. (Sofia, Bulgaria) under the terms of a contract with the European Commission Joint Research Center. This is capable of making structure-based predictions for several toxicological endpoints including skin sensiti	Toxic Hazard Estimation by decision tree approach: http://toxtree.sourceforge.net/
VirtualToxLab™	The VirtualToxLab™ is an *in silico* tool for predicting the toxic potential (endocrine and metabolic disruption, interference with the hERG ion channel) binding affinity towards (currently) 16 target proteins: AhR, AR, ER, hERG, GR, LXR, MR, PPAR, TR, CYP	VTV Molecular Modelling: http://www.chemie.unibas.ch/~vbc/molmod/virtualtox/index.html

particular toxicity endpoint. In this approach, the steric and electronic environment surrounding a given structural alert fragment are easily ignored but these factors can either diminish or enhance its toxicity. Moreover, this method is only designed for highlighting positives (i.e. if no alerts are present, it does not mean the chemical will not have any toxic effects), and also the role of these small fragments towards many tox endpoints is not well understood [44]. Where possible, modellers should employ several predictive models for a single endpoint to produce better consensus predictions [43]. For example in the case of predicting an Ames test outcome, integration of human derived structural alerts with artificial intelligence systems for Ames in a consensus modelling manner has been shown to provide advantages over that of a single model (User's Guide for T.E.S.T: http://www.epa.gov/nrmrl/std/cppb/qsar/testuserguide.pdf).

It is also possible that some models might be better at predicting some subclasses of chemicals; it is therefore also possible to create substructure-localised consensus models by taking into account the strengths of each model for a particular substructural class (User's Guide for T.E.S.T: http://www.epa.gov/nrmrl/std/cppb/qsar/testuserguide.pdf) [45,46]. It should be noted that consensus modelling might not offer any advantages in cases where a strong model is integrated with many weaker models, in cases like these consensus prediction might even offer low predictivity over single model due to noise addition from the weaker models.

Before model building, modellers need to have a clear picture about where and how models will be used. It is important to highlight the benefits and limitations of each model so that users are able to understand when and why models might fail. Users of the models also need to be aware of each step in the model building workflow, including which data was used in the model building exercise, which descriptors were used and what is the role of each descriptor. For example, if a model is built on simple 2D descriptors or fragments then it might not be able to highlight differences in toxicity of stereo-isomers or even similar chemicals in 2D representations. It is often seen that the interpretation of the success of models depends on their use and the expectations of the user [47]. Therefore it is not surprising that for the same data and same model, one user might find value in the model, whereas for another user the same model might be of little value. For users it is therefore important to know the building process of models and expectations from it especially when these might be promoted for use in a regulatory setting.

SAR, QSAR and read-across are data mining approaches, which involve analysis of the structural features of sets of chemicals to generate rules that enable users to predict outcomes for new chemicals. The use of these approaches in a more systematic and integrated workflow is suggested in Fig. 1. For any given target chemical, one of the first steps should be to look for what is known already for that chemical by use of text and data extraction techniques. If there are gaps in the data, the next step should be to fill these gaps using close analogues. Here, the strength of available evidence should be judged on a case-by-case basis. For example, in a case where there are many close analogues with similar toxicology profiles, confidence will be much higher as compared with a case where there are few close analogues with a mixed toxicology profile. QSAR might be used as the final step if there is no toxicology data available either on exact or close analogues, but all the points about applicability domain and confidence in predictions discussed above should be taken into account.

Limitations of (Q)SAR approaches

Overall it also needs to be noted that these individual (Q)SAR *in silico* toxicological methods are hazard identification methods and in most of cases they do not take dose and exposure into account unless a exposure–response relationship has been studied. Therefore, in general, these will not predict toxicity in isolation, but provide useful supplementary information for the overall risk assessment process. For example, the aromatic nitro group is a

Please cite this article in press as: S.. Modi, et al., The value of *in silico* chemistry in the safety assessment of chemicals in the consumer goods and pharmaceutical industries, Drug Discov Today (2011), doi:10.1016/j.drudis.2011.10.022

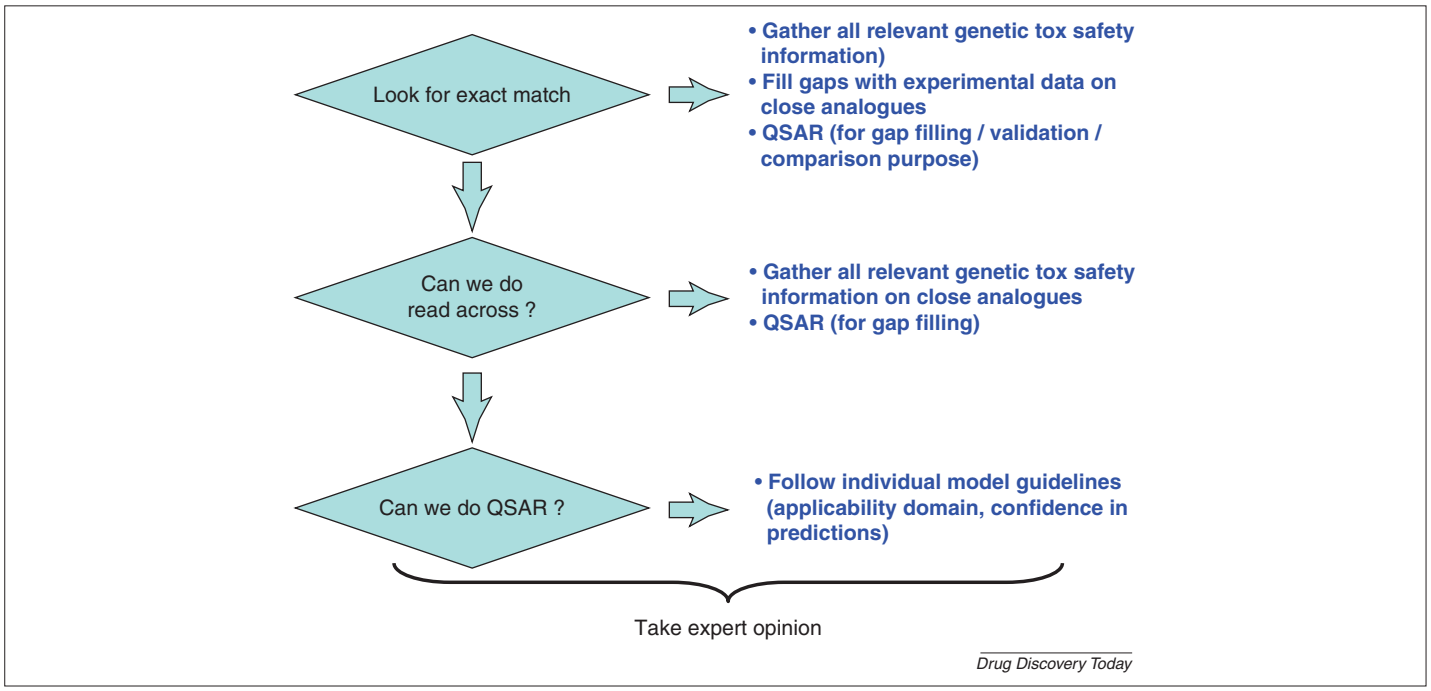

FIGURE 1

A suggested integrated workflow for combined use of data-mining, (Q)SAR and read-across methods. Abbreviation: (Q)SAR: quantitative structure activity relationships.

well known fragment that triggers a structural alert for carcinogenicity, but if a chemical containing this fragment has very low exposure or bioavailability, it is questionable whether this prediction will be realised. Therefore, whenever possible, internal exposure (i.e. the amount taken up and distributed as free plasma concentration within an organism) should be taken into account by using either *in silico* or *in vitro* ADME data. Ideally, results of predictions should be combined with other evidence and data for consideration for the risk assessments.

Another limitation for most of the current (Q)SAR approaches is that they mostly do not consider metabolites of the parent chemicals, this is usually true for endpoints like hERG, skin sensitisation, among others [34,35]. When properties are calculated for prediction of a toxicology endpoint, it is usually conducted on a parent chemical structure, whereas it could be a metabolite which is responsible for the toxicity. However, clearance by metabolism can also have an important role in the actual exposure to a given chemical. Safety guidance has been produced, with triggers for concern based on the abundance of metabolites relative to total material or relative to parent material levels [48,49]. Others have proposed a strategy in which absolute exposure to metabolites (rather than a relative comparison with parent or total drug-related material) in humans triggers further consideration of metabolite safety [50]. It is possible to predict metabolites using several computational approaches (METASITE [51], METEOR [52], Metaprint2D [53], among others). However, it should be noted that the majority of these methods simply predict qualitatively the metabolites that could be formed and do not estimate the probability or amount of each metabolite being formed.

By definition, all of these *in silico* models for toxicity are simulations of reality; the interplay of complex physiological processes that lead to toxicity presents a real challenge for creating reliable predictive models [5]. But greater confidence can be obtained if all other possible information, including limitations of models, applicability domain, metabolites and exposure are combined together for the risk assessments [42,47].

Concluding remarks

Experimentalists generally prefer to generate 'wet' data on all the chemicals, irrespective of what the odds of success might be in terms of late-stage failures. With a lack of confidence in the accuracy of a predictive model, *in silico* tools tend to carry little weight in a risk assessment. As a result, the level of acceptance of predictions by users outside the computational chemistry and modelling groups tends to be low. There is need to be transparent wherever possible and modelling methods need to be chosen carefully, including confidence factors and reasoning behind each of the predictions. Every modeller must be encouraged to promote their use in the context of realistic expectations of these tools. Modellers need to do more than just generate large numbers of data points, they need to work within multi-disciplined program teams to provide the support needed for the decision-making process in a project. As more and more validated case examples are passed through these different *in silico* approaches it might help in gaining confidence and understanding of these tools [19,54,55].

Keeping all the above in mind, it is not surprising that the acceptance of these predictive tools continues to be difficult, but if the existing issues can be appropriately addressed, it might eventually be well worth the effort. In addition there are commercial and consumer pressure to find and use alternatives for less environmental impact and also less animal testing. Therefore an

Please cite this article in press as: S.. Modi, et al., The value of *in silico* chemistry in the safety assessment of chemicals in the consumer goods and pharmaceutical industries, Drug Discov Today (2011), doi:10.1016/j.drudis.2011.10.022

opportunity exists for these *in silico* methods to demonstrate real value as part of a suite of alternative technologies and gain public support of this area of science.

There is also need for computational chemistry tools to align with other information sources (e.g. from systems biology, hazard, metabolites and exposure) to develop real or virtual models of tissues, organs and physiological processes that could be used for the toxicological assessments. For toxicity assessments, information from other sources could be applied as part of a tiered system along with predictions from computational tools. As suggested in Fig. 2, the first tier could be used for alerts, (Q)SAR and read-across methods could be used for hazard identification. These tools might represent a fast method and filter to enrich a biological screen with desired ADMET profile. Indeed, many major pharmaceutical companies have already adopted virtual screening methodologies to complement *in vitro* HTS methods [56,57].

It is also important for these tools to move beyond hazard rankings and possibly move towards estimation of *in vivo* responses based on *in vitro* or *in silico* data. As discussed above predictions based just on *in silico* for prioritising chemicals might over- or under-estimate the potential risk of these chemicals owing to differences in bioavailability, clearance and exposure. Physiologically based pharmacokinetic (PBPK) models can take *in vitro* and *in silico* data inputs and can predict concentration versus time profiles. It has been shown in the past that PBPK models are superior to other more empirical methods for interspecies scaling and prediction of human pharmacokinetics [58]. This approach could be the second tier described in Fig. 2 integrating both dosimetry and human exposure information with the *in silico*, high-throughput, toxicity screening data to provide a better safety risk assessment. A recent article [59] has shown the combined use of experimental assays, computational tools, and exposure assessment by performing analysis on a subset of 35 ToxCast chemicals.

Understanding the different mechanisms of how chemicals can affect biological structures, processes and pathways and thus can impact on physiological response is an important aspect of toxicology. This knowledge can help to predict the toxicity of chemicals, and it is also possible to plan ways to prevent exposure to toxic compounds and develop ways to antagonise the effects of the toxins. Both the complexity in the biological response and the lack of public availability for mechanistic data that can be modelled to relevant structural information are reasons that *in silico* approaches to date have had limited success in delivering *in vivo* relevant predications. In view of this, the European commission Seventh Framework Programme (FP7) research joint technology programmes, the Innovative Medicines Initiative (IMI: http://www.imi.europa.eu/) funded with the European Federation of Pharmaceutical Industries and Associations (EFPIA) and the Safety Evaluation Ultimately Replacing Animal Testing (SEURAT1) project (Colipa: http://www.colipa.eu/news-a-events.html) funded

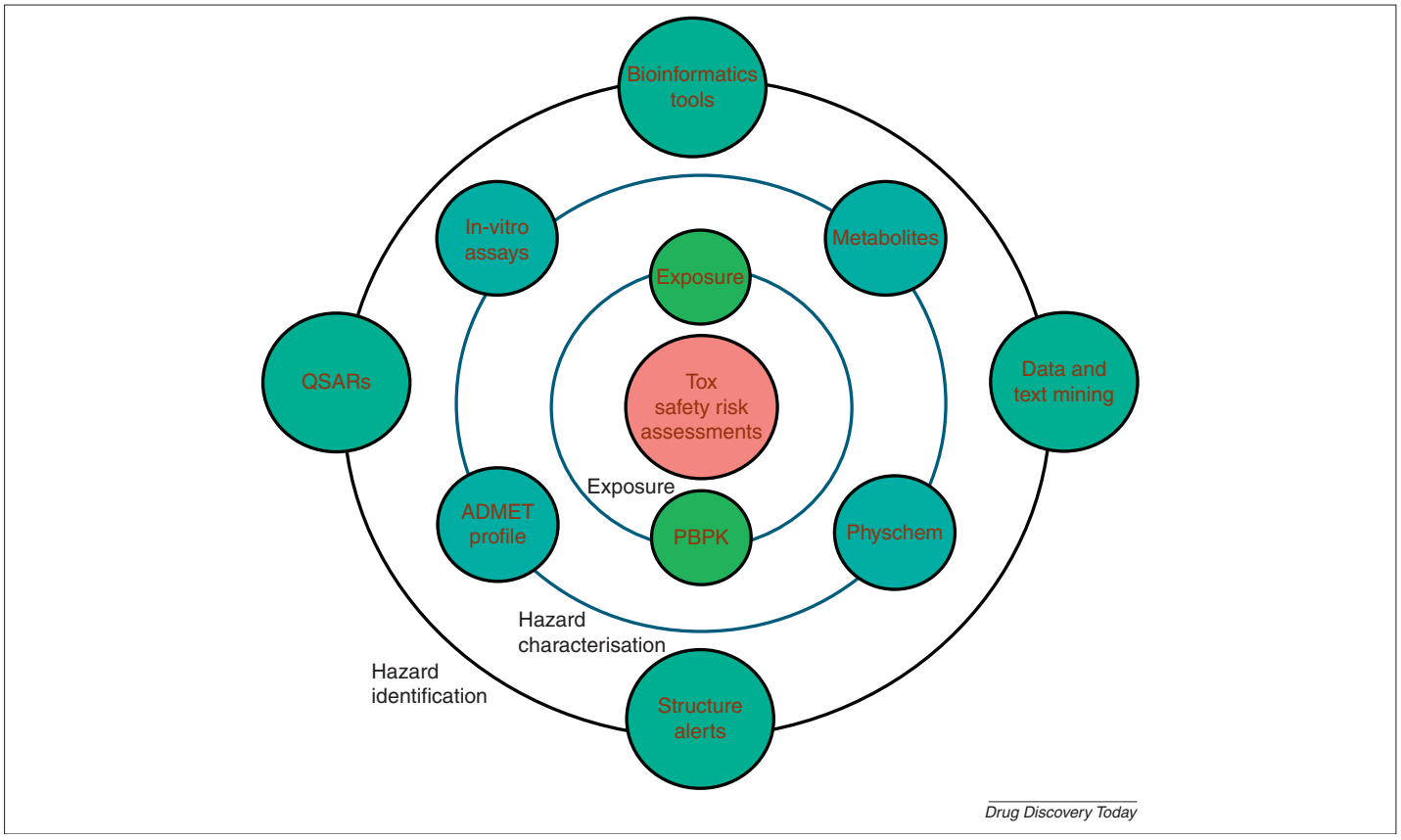

FIGURE 2

List of some of the parameters input which have important roles in different tiers (hazard identification, characterisation and exposure) for the toxicity safety assessments. Abbreviations: ADMET: absorption distribution metabolism elimination and toxicity; PBPK: physiologically based pharmacokinetic; (Q)SAR: quantitative structure activity relationships.

Please cite this article in press as: S.. Modi, et al., The value of *in silico* chemistry in the safety assessment of chemicals in the consumer goods and pharmaceutical industries, Drug Discov Today (2011), doi:10.1016/j.drudis.2011.10.022

with European Cosmetic, Toiletry and Perfumery Association (Colipa), have undertaken several projects, such as eTox, Expert Systems for *in silico* Toxicity, and COSMOS. These programmes although focussed on different end goals of improving the efficiency of drug development (IMI) and non-animal alternatives to repeat dose toxicity respectively (SEURAT-1) they both aim to deliver improvements in the modelling of early prediction of *in vivo* human toxicity based on information and data available during early stages of the innovation pipeline. Furthermore mechanistic understanding of human toxicity forms [60] a central component of the National Research Council of the National Academies (NRC) vision and roadmap as described in 'Toxicity Testing in 21st Century (TT21C): A Vision and a Strategy' [61]. This vision is summarised as follows: 'Advances in toxicogenomics, bioinformatics, systems biology, epigenetics and computational toxicology could transform toxicity testing from a system based on whole-animal testing to one founded primarily on *in vitro* methods that evaluate changes in biologic processes using cells, cell lines, or cellular components, preferably of human origin.' The national toxicology programme (NTP) HTS initiative and the U.S. Environmental Protection Agency (EPA) Toxcast program [59,62] are two efforts that aim to utilise the technological advances in molecular biology and computational science. These aim to identify toxicological testing screens for mechanistic targets active within cellular pathways considered crucial to adverse health effects, such as carcinogenicity, reproductive and developmental toxicity, genotoxicity, neurotoxicity, and immunotoxicity in humans. However, one of the main outstanding issues remains that knowing that changes in biological processes and perturbation in pathways occur might not be good enough: we need to define how much change is really required for a chemical to cause an adverse effect. These *in silico* techniques and preclinical testing (*in vitro*) serve a fundamental role in characterising of the potential risks associated with chemicals. However, serious and sometimes rare and unexpected adverse events might be observed in clinical trials or post-approval, suggesting that crucial gaps exist in our understanding of the relationship between patient response and preclinical toxicology findings [12,13]. For example, non-clinical safety assessment are often conducted in normal healthy test systems and tends to be exposure-based; it does not attempt to evaluate the possible risk of rare or idiosyncratic responses that might arise from potential interactions with the presence or progression of disease or the genetic background or other exposures of patients and consumers. Therefore to improve the predictions of chemical safety several programmes mentioned above (e.g. TT21C, COSMOS, ToxCast, among others) have been started to gain a better understanding of toxicity mechanisms by evaluating safety assessment data at multiple levels of biological organisation, including genes, proteins, pathways and cell/organ function. There is also need to develop computer models of cells, organs and system and develop clinical trial simulation models that can reveal interactions between drug or device effects, patient characteristics, and disease variables influencing outcomes [63]. As more and more case examples are passed through these different *in silico* approaches it might help to gain more understanding on how these should be linked together.

There is still a need for new methods to rapidly and accurately determine the toxic potential of both drug molecules and molecules contained in consumer product goods. *In silico* toxicology models, such as those discussed in this article, fit many of these criteria, and have seen widespread use in drug discovery applications. It should also be acknowledged that *in silico* tools have been in existence for a relatively short time compared with *in vitro* or *in vivo* methods, and have only taken up a pace in past 10 years. As these tools become increasingly user-friendly and transparent, and as more examples of successful applications are shown, it seems highly probable that *in silico* approaches might evolve rapidly. Meanwhile, modellers need to ensure that these tools are used appropriately and expectations of users are not raised above what a model can deliver. Rather than just a static tool sitting on its own in one corner it needs to be integrated with safety assessments and discovery programmes. By working alongside drug discovery programmes and with constant validation of these *in silico* methods against *in vitro* and *in vivo* data, predictive chemistry can become an increasingly important part of the decision-making process. Where possible as much information as possible needs to be integrated as required for individual toxicity assessments and used as part of a weight of evidence approach to calculate risk to the consumer. At present we have most of the individual components for building a platform for virtual models of tissues, organs and physiological processes that could be used for the toxicological assessments, but one of the main challenges will not be to see how these different components from different disciplines fit together. As discussed above, as more and more examples and case studies are pushed through we might become better and better in gaining knowledge and learn from these information rich technologies.

References

1 Eddershaw, P.J. *et al.* (2000) ADME/PK as part of a rational approach to drug discovery. *Drug Discov. Today* 5, 409–414
2 Smith, D.A. and Waterbeemd, H. (1999) Pharmacokinetics and metabolism in early drug discovery. *Curr. Opin. Chem. Biol.* 3, 373–378
3 Wang, J. and Urban, L. (2004) The impact of early ADME profiling on drug discovery and development strategy. *Drug Discov. World Spring* 5, 73–86
4 Schumann, R. (2002) The seventh amendment to the cosmetics directive: what does DG enterprise want from ECVAM? *Altern. Lab. Anim.* 30 (Suppl. 2), 213–214
5 Kimber, I. *et al.* (2011) Computational chemistry, systems biology and toxicology. Harnessing the chemistry of life: revolutionizing toxicology. A commentary. *J. Appl. Toxicol.* 31, 206–209
6 Vedani, A. and Smiesko, M. (2009) In silico toxicology in drug discovery – concepts based on three-dimensional models. *Altern. Lab. Anim.* 37, 477–496
7 Modi, S. (2004) Positioning ADMET in silico tools in drug discovery. *Drug Discov. Today* 9, 14–15
8 Waterbeemd, H. and Gifford, E. (2003) ADMET in silico modelling: towards prediction paradise? *Nat. Drug Discov. Rev.* 2, 192–204
9 Dickins, M. and Modi, S. (2002) Importance of predictive ADME simulations. *Drug Discov. Today* 7, 755–756
10 Cronin, M. *et al.* (2003) Use of QSARs in international decision-making frameworks to predict health effects of chemical substances. *Environ. Health Perspect.* 111, 1391–1401
11 Netzeva, T.I. *et al.* (2005) Current status of methods for defining the applicability domain of (quantitative) structure–activity relationships. The report and recommendations of ECVAM Workshop 52. *Altern. Lab. Anim.* 33, 155–173
12 Hartung, T. (2008) Food for thought on animal tests. *ALTEX* 25, 3–9
13 Hartung, T. (2009) Toxicology for the 21st century. *Nature* 460, 208–212

Please cite this article in press as: S.. Modi, et al., The value of *in silico* chemistry in the safety assessment of chemicals in the consumer goods and pharmaceutical industries, Drug Discov Today (2011), doi:10.1016/j.drudis.2011.10.022

14 Clark, D.E. (2007) In silico ADMET tools: the dawn of a new generation? *Expert Opin. Drug Discov.* 2, 1423–1429
15 Williams, A. *et al.* (2010) Free online resources enabling crowd-sourced drug discovery. *Drug Discov. World* 10, 33–39
16 Fourches, D. *et al.* (2010) Trust, but verify: on the importance of chemical structure curation in cheminformatics and QSAR modeling research. *J. Chem. Inf. Model.* 50, 1189–1204
17 Kazius, J. *et al.* (2005) Derivation and validation of toxicophores for mutagenicity prediction. *J. Med. Chem.* 48, 312–320
18 Schultz, T.W. (2007) Verification of the structural alerts for Michael acceptors. *Chem. Res. Toxicol.* 20, 1359–1363
19 Valerio, L.G. (2009) In silico toxicology for the pharmaceutical sciences. *Toxicol. Appl. Pharmacol.* 241, 356–370
20 Judson, P.N. (2006) Using computer reasoning about qualitative and quantitative information to predict metabolism and toxicity. In *Pharmacokinetic Profiling in Drug Research: Biological, Physicochemical, and Computational Strategies* (Testa, B. *et al.* eds), pp. 183–215, Wiley
21 Schultz, T.W. *et al.* (2003) The present status of QSAR in toxicology. *J. Mol. Struct. Theochem.* 622, 23–38
22 Cariello, N.F. *et al.* (2002) Comparison of the computer programs DEREK and TOPKAT to predict bacterial mutagenicity. *Mutagenesis* 17, 321–329
23 Benfenati, E. *et al.* (2009) Predictive models for carcinogenicity and mutagenicity: frameworks, state-of-the-art, and perspectives. *J. Environ. Sci. Health C: Environ. Carcinog. Ecotoxicol. Rev.* 27, 57–90
24 Hansen, K. *et al.* (2009) Benchmark data set for in silico prediction of Ames mutagenicity. *J. Chem. Inf. Model.* 49, 2077–2081
25 Mazzatorta, P. *et al.* (2008) Modeling oral rat chronic toxicity. *J. Chem. Inf. Model.* 48, 1949–1954
26 Bercu, J. *et al.* (2010) In silico approaches to predicting cancer potency for risk assessment of genotoxic impurities in drug substances. *Regul. Toxicol. Pharmacol.* 57, 300–306
27 Vink, S.R. *et al.* (2010) Use of read-across and tiered exposure assessment in risk assessment under REACH – a case study on a phase-in substance. *Regul. Toxicol. Pharmacol.* 58, 64–71
28 Yang, C. *et al.* (2008) Understanding genetic toxicity through data mining: the process of building knowledge by integrating multiple genetic toxicity databases. *Toxicol. Mech. Methods* 18, 277–295
29 Bajorath, J. (2002) Virtual screening in drug discovery: methods, expectations and reality. *Curr. Drug Discov.* 2, 24–28
30 Nonoyama, T. and Fukuda, R. (2008) Drug-induced phospholipidosis – pathological aspects and its prediction. *J. Toxicol. Pathol.* 21, 9–24
31 Crivori, P. and Poggesi, I. (2006) Computational approaches for predicting CYP-related metabolism properties in the screening of new drugs. *Eur. J. Med. Chem.* 41, 795–808
32 Li, H. *et al.* (2008) Considerations and recent advances in QSAR models for cytochrome P450-mediated drug metabolism prediction. *J. Comput. Aided Mol. Des.* 22, 843–855
33 Michielan, L. and Moro, S. (2010) Pharmaceutical perspectives of nonlinear QSAR strategies. *J. Chem. Inf. Model.* 50, 961–978
34 Li, Q. *et al.* (2008) hERG classification model based on a combination of support vector machine method and GRIND descriptors. *Mol. Pharm.* 5, 117–127
35 Gavaghan, C.L. *et al.* (2007) Development, interpretation and temporal evaluation of a global QSAR of hERG electrophysiology screening data. *J. Comput. Aided Mol. Des.* 21, 189–206
36 Jaworska, J. *et al.* (2005) QSAR applicability domain estimation by projection of the training set descriptor space: a review. *Altern. Lab. Anim.* 33, 445–459
37 Stanforth, R.W. *et al.* (2007) A measure of domain of applicability for QSAR modelling based on intelligent K-means clustering. *QSAR Comb. Sci.* 26, 837–846
38 Dragos, H. *et al.* (2009) Predicting the predictability: a unified approach to the applicability domain problem of QSAR models. *J. Chem. Inf. Comput. Sci.* 49, 1762–1777
39 Bentzien, J. *et al.* (2010) An in silico method for predicting ames activities of primary aromatic amines by calculating the stabilities of nitrenium ions. *J. Chem. Inf. Model.* 50, 274–297
40 Kar, S. and Roy, K. (2010) QSAR modeling of toxicity of diverse organic chemicals to Daphnia magna using 2D and 3D descriptors. *J. Hazard. Mater.* 177, 344–351
41 Eriksson, L. *et al.* (2003) Methods for reliability and uncertainty assessment and for applicability evaluations of classification- and regression-based QSARs. *Environ. Health Perspect.* 111, 1361–1367
42 Tropsha, A. *et al.* (2003) The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models. *QSAR Comb. Sci.* 22, 69–79
43 Abshear, T. *et al.* (2006) A model validation and consensus building environment. *SAR QSAR Environ. Res.* 17, 311–321
44 Mathews, E.J. (2004) Assessment of the health effects of chemicals in humans: I. QSAR estimation of the maximum recommended therapeutic dose (MRTD) and no effect level (NOEL) of organic chemicals based on clinical trial data. *Curr. Drug Discov. Technol.* 1, 61–76
45 Baurin, N. *et al.* (2004) 2D QSAR consensus prediction for high-throughput virtual screening. An application to COX-2 inhibitionmodeling and screening of the NCI database. *J. Chem. Inf. Comput. Sci.* 44, 276–285
46 Guha, R. and Schürer, S.C. (2008) Utilizing high throughput screening data for predictive toxicology models: protocols and application to MLSCN assays. *J. Comput. Aided Mol.* 22, 367–384
47 Stouch, T.R. *et al.* (2003) In silico ADME/Tox: why models fail. *J. Comput. Aided Mol. Des.* 17, 83–92
48 Baillie, T.A. *et al.* (2002) Drug metabolites in safety testing. *Toxicol. Appl. Pharmacol.* 182, 188–196
49 United States Food and Drug Administration: Center for Drug Evaluation and, Research, (2008) *Guidance for Industry Safety Testing of Drug Metabolites.*
50 Smith, D.A. and Obach, R.S. (2005) Seeing through the mist: abundance versus percentage. Commentary on metabolites in safety testing. *Drug Metab. Dispos.* 33, 1409–1417
51 Cruciani, G. *et al.* (2009) Integrating crystallography into early metabolism studies. In *From Molecules to Medicine* (Sussman, J.L. and Spadon, P., eds), pp. 63–77, Springer Science
52 Button, W.G. *et al.* (2003) Using absolute and relative reasoning in the prediction of the potential metabolism of xenobiotics. *J. Chem. Inf. Comput. Sci.* 43, 1371–1377
53 Boyer, S. (2007) Reaction site mapping of xenobiotic biotransformations. *J. Chem. Inf. Model.* 47, 583–590
54 Cronin, M.T.D. (2009) In silico toxicology challenges for pharmaceuticals: complacency or controversy? *Altern. Lab. Anim.* 37, 453–456
55 Raunio, H. (2011) In silico toxicology – non-testing methods. *Front. Pharmacol.* 2, 1–8
56 Modi, S. (2003) Computational approaches to the understanding of ADMET properties and problems. *Drug Discov. Today* 8, 621–623
57 Fostel, J. (2005) Predictive ADME-Tox. *Expert Opin. Drug Metabol. Toxicol.* 1, 565–570
58 Parrott, N. *et al.* (2005) Application of full physiological models for pharmaceutical drug candidate selection and extrapolation of pharmacokinetics to man. *Basic Clin. Pharmacol. Toxicol.* 96, 193–199
59 Rotroff, D. (2010) Xenobiotic-metabolizing enzyme and transporter gene expression in primary cultures of human hepatocytes modulated by toxcast chemicals. *J. Toxicol. Environ. Health B: Crit. Rev.* 13, 1329–1346
60 Schmidt, C.W. (2009) TOX 21: new dimensions of toxicity testing. *Environ. Health Perspect.* 117, A348–A353
61 Andersen, M. and Krewski, D. (2009) Toxicity testing in the 21st century: bringing the vision to life. *Tox. Sci.* 107, 324–330
62 Knudsen, T.B. *et al.* (2011) Activity profiles of 309 ToxCast™ chemicals evaluated across 292 biochemical targets. *Toxicology* 282, 1–15
63 Shah, I. *et al.* (2010) Virtual tissues in toxicology. *J. Toxicol. Environ. Health B: Crit. Rev.* 13, 314–328

Please cite this article in press as: S.. Modi, et al., The value of *in silico* chemistry in the safety assessment of chemicals in the consumer goods and pharmaceutical industries, Drug Discov Today (2011), doi:10.1016/j.drudis.2011.10.022