Bridging Disciplines: The Role of Computational Approaches in Toxicology and Risk Assessment

[1]Shivam D Kulkarni, [2]Kanchan S Tone [3]Akash V Gangane, [4]Vishwajit R Ghodke, [5]Tejas S Patil, [6]Mohit D Chamkure [7] Mayur D Patil*
[1]Dayanand college of pharmacy latur MS INDIA, [2]Prerna institute of pharmacy parbhani, Parbhani
[3-4]Dayanand college of pharmacy latur MS INDIA, [5]Channabasweshwar Pharmacy College, Latur,
[6]MDA College of Pharmacy Kolpa [7]Dayanand college of pharmacy latur MS INDIA,
[1-6]Department of Pharmaceutics

***Abstract*—[In the spectrum of industrial sectors pharmaceuticals chemicals personal care products food additives and their associated regulatory agencies, there is a need to develop robust and reliable methods to reduce or test. It is generally recognized that no single alternative method will be able to provide a one-to-one replacement for assays based on more complex toxicological endpoints Hence information from a combination of techniques is required. A greater understanding of the time and concern Across traction-dependent mechanisms, underlying the interactions between chemicals and biological systems and the sequence of events that can lead to apical effects will help to move forward the science of reducing and replacing animal experiments. In silico modeling in vitro, assays high-throughput screening organ-on-a-chip technology omics and mathematical biology can provide complementary information to develop a complete picture of the potential response of an organism to a chemical stressor. Adverse outcome pathways (AOPs) and systems biology frameworks enable relevant information from diverse sources to be logically integrated. [1]**

I. INTRODUCTION

The agenda for change has been mandated with a global drive towards reducing refining or replacing animal tests with non-animal alternatives. Legislative changes and commercial and ethical pressures have demotivated the pursuit of alternatives to the traditional in vivo assays used in product development and safety assessment. However tangible progress varies significantly between sectors and geographic regions. For example cosmetic products or their ingredients to be marketed. Within the European Union (EU) can no longer be tested on animals whereas the registration of new therapeutic entities (NTEs) is contingent upon safety efficacy and protocols being established in animal models.[2]

To develop products that are safe for humans and animals (including environmental species) there is a need to understand the potential effects of chemicals on a wide range of organisms and how this can be affected by factors such as developmental stage health status or individual genetic composition. The ability of a chemical to elicit an effect is determined by its concentration-time profile (internal exposure) at a relevant site and its inactivity (toxicity). Data may be available for certain species under specific exposure scenarios for some chemicals. However, there are no chemicals for which comprehensive data are available for all target and non-target species that. [3]

II. OVERVIEW OF IN SILICO TOOLS AND THEIR APPLICATIONS.

In silico tools use existing data or information derived from molecular structure to make predictors regarding absorption distribution metabolism and excretion (ADME process that determines internal exposure) and biological activity (toxicity or hazard) of a chemical.

• Methods range from simple (Quantitative) structure-activity (property) relationships to more complex models.

Integrating information from different sources enables more accurate predictions of chemical activity used in product development and safety assessment.

Knowledge of the key terminology vend in this area of science is important to understand this review here commonly used abbreviations and key terms that are used herein have been summarized For cample, the term new approach methodology (NAM) applies to arty of the non-annual alternatives

that can be used alone or in combination to provide information for safety assessment. Integrated approaches to testing and assessment (GATAs)combine information from a range of sources to determine if there is sufficient knowledge on which to make safety-based decisions or direct future experiments to fill knowledge gaps. "Defined approaches (DAs) use a fixed data interpretation procedure (DIP) to interpret information from a defined set of sources to assist decision making It is the combination of data on absorption The power of in silico models is that the predictions are derived entirely from the structures of the chemicals of interest. The fundamental tenet of these models is that the intrinsic properties, potential interactions, and ultimate effects of a chemical are encoded within its molecular structure, understanding this enables (quantitative) structure-activity relationship ((Q)SAR) or (quantitative) structure-property relationship ((Q)SPR) models to be developed. Similar chemicals are expected to elicit similar effects.

Hence, knowledge of one chemical (or a group of chemicals) can be used to predict the characteristics of similar chemicals. Limitations to this approach recognized for example, the level of sophistication by which the molecule can be described at a structural level and the role of mitigating factors (e.g. structural features that may attenuate or intensify a response) can alter activity. The similarity-paradox refers to the problem of chemicals that are ostensibly similar but exhibit markedly different Activity profiles. Activity cliffs occur when a smooth relationship between structure and activity abruptly ceases and a small change in structure leads to an extreme change in biological response. Activity cliffs may provide an opportunity for medicinal chemists to develop new leads but they are problematic in developing QSARs. Notwithstanding in silico models have been used for a multitude of applications from predicting the toxicity of pollutants or agrochemicals to environmental species, to optimizing drug candidates. The history of the development of in silico tools has been summarized recently. Thousands of [1]

The purpose of this review is neither to recount the complete history of the fields nor to catalog large numbers of available models or software applications. Rather its purpose is to provide an introduction to the breadth of tools available and the underlying theory and applications of these tools for those new to the area of in silico prediction. The range of techniques and their applications for example prediction of intrinsic activity (hazard) or internal exposure (ADME properties) are summarized as key concepts in Figure 1 and explained in detail below Examples are also given regarding the use of these tools across different sectors e.g. drug development in the pharmaceutical industry, safety assessment in the personal care product and food industries and environmental toxicity prediction. An explanation of the theories that underpin the key methods, and how model reliability may be evaluated is also presented with additional resources (eg. exemplar software and comprehensive reviews of individual methods) being signposted within the relevant section.

III APPLICATIONS OF IN SILICO MODELS

As the number of in situ tools has expanded, so too has their application across different industrial and regulatory sectors. This is advantageous, as new information on the development or application of tools in one sector can be leveraged by another. This cross-disciplinary sharing of ideas and practice enables more rapid advancement acceptance and uptake of new in silico methods. Historically, in silico models have been widely used for predicting the invicity of chemicals to environmental species particularly fish aquatic invertebrates, algae, and more recently bees However there are multiple examples of their application within medicinal chemistry in the design of bioactive chemicals, predictive toxicology, and safety assessment. In the pharmaceutical industry models have been used extensively to maximize the efficiency of the drug development process to ensure that only those candidates likely to be successful are taken forward to the animal testing stages and to avoid late-stage attrition or post-marketing withdrawal The personal care product (agro)chemical and food industries also use a range of in silico tools in product development and as a result there is now increased recognition of the potential of in silico tools to provide information for regulatory submissions to meet legislative demands The EU regulation concerning the Registration Evaluation Authorization and restriction of Chemicals (REACH)which came into force in June 2007 aims to protect humans and the environment from the adverse effects of the use of chemicals. [5]

The REACH regulation specifically promotes the use of in silico prediction (e.g. QSAR and read-across methods) as alternatives to animal testing, providing that the results are derived from a (Q)SAR. Model for which scientific validity has been established, the substance falls within the applicability domain of the (Q)SAR model the results are adequate for the purpose (e.g. classification or labeling) and adequate and reliable documentation of the applied method is provided. ECHA's 4th Report on the Use of Alternatives to Testing on Animals for the REACH Regulations confirms that results from alternative methods continue to be used over and above new animal tests in dossiers submitted for REACH. Read-across is the most common alternative strategy but use is also made of QSARs waiving and integrated testing strategies (ITS).

Since 2013 the Cosmetics Regulation (Regulation (EC) No 1223/2009) has banned the testing of cosmetic ingredients and products on animals and has prohibited the marketing of cosmetics for which the ingredients or products were tested on animals since the introduction of the ban. The 10th revision of the Notes of Guidance for the testing of cosmetic ingredients and their safety evaluation (from the European Commission's Scientific Committee on Consumer Safety (SCCS)) similarly promotes the use offing silico models stipulating that for safety evaluation of cosmetic ingredients all available scientific data are con-side red including results from (Q)SARs chemical categories grouping read across and physiologically-based kinetic (PBK) modeling. For the safety assessment of food and food ingredients a stepwise roadmap for evaluation that draws upon information from in silico

Models including QSAR and read-across have been proposed. In the USA the Frank R Lautenberg Chemical Safety for the 21st Century Act of Congress (2016) enshrines into US law that animal studies should be reduced or replaced as much as practicable. A cross-sector partners forum (organized by the European Partnership for Alternatives to Animal Testing (EPAA)) resulted in a report on the use of read-across by the pharmaceutical cosmetics (personal care product)chemical agrochemical food and fragrance industries and their associated regulatory organizations. The report identified cross-industry synergies in approaches and highlighted the need to incorporate toxic kinetic information in read-across. This evidences the increasing use of in silico tools across all sectors highlighting the extensive economic and ethical contribution of this area of science. [8]

Databases

The rate of acquisition of scientific knowledge is expanding more rapidly now than at any other time. To maximize the value of this new information there is a significant need for much of it to be made available in the public domain. This has led to a rapid expansion of databases often freely accessible that can provide a wealth of information on millions of chemicals. Whether or not a database itself can strictly be defined as an in silico tool is debatable. However, the searching strategies incorporated within modern databases (eg algorithms to identify similar chemists or the capacity to modify and combine search parameters) are certainly some of the most widely used tools in finding data on chemicals to develop and/or evaluate models. Often chemicals are characterized as being data-rich (having a high volume of relevant data readily available) or data-poor (having little or no relevant data available).

The interrogation of existing databases is a vital first step in determining the potential effects of a chemical inappropriate data are already available then this obviates the need for testing or generating predictions. Generally, it is better to use an experimental rather than predicted value (unless there are known problems with the experiment) where multiple values are available judgment or consideration of data quality must be applied as discussed in the next section. If information is not available for the chemical of interest it may be possible to predict rational use of information available for other chemicals this practice underpins in silico modeling tools such as (Q)SAR and read-across. Identifying the existing data also highlights where there are knowledge gaps and therefore can help to prioritize future testing strategies ensuring that the maximum information is obtained from those chemicals that are selected for testing. Databases are usually searchable by using a range of chemical identifiers such as

- name
- Simplified Molecular Input Line Entry Systems (SMILES) string
- hashed code derived from the International Chemical Identifier (i.e. InChIKey)

- Registry number (e Chemical Abstracts (CAS) se European Inventory of Existing Chemical Substances (EINECS) number)

Paramount importance is ensuring that the data obtained (activity toxicity hazard data or ADME values) have been correctly and unambiguously assigned to the correct chemical structure. With increasing automation it is easy for errors in chemical names or structures to be propagated in databases or literature collation. High-quality databases often report the methods used to assess data record accuracy and may have standard protocols for error reporting and fixing. Cross-checking that at least two and ideally three identifiers give consistent results can be performed to check consistency in structure identification where inconsistencies are identified primary literature may need to be consulted. Some databases offer the capability to search for chemicals that are similar to the chemical of interest by using chemical fingerprints (vide infra) and/or physicochemical properties. A recent comprehensive review of over 900 databases was identified and characterized in terms of the type of information available as well as their public or commercial accessibility interoperability search criteria etc. The categories for the types of database considered (with the associated databases given in parentheses) were biological (268) drug discovery (157) clinical trials (116)chemistry (80) omics (60) toxicology (57) protein-protein interactions (54) alternative methods (39) ADME (38) pathways (38) environmental exposure (30)nonmaterial's toxicity and patents (9). Of the hundreds of databases available some representative examples of freely accessible databases are to indicate the nature and scope of these resources. PubChem is one of the most comprehensive sources of chemical information. It can be searched by using name synonyms molecular formula structure SMILES InChIKey or registry number. It is also possible to search for chemicals that are similar concerning 2-

D fingerprint or physicochemical properties.

The type of information available is divided into approximately 20 major categories depending on the nature of the chemical) and each major category expands into multiple subcategories providing information on for example identifiers and physical properties using pharmacology safety/hazarıddata and toxicity data references. Similarly, Chemspideris another comprehensive resource with information on identifiers' physical properties and chemical properties (experimental and/or predicted values) with links to predictions from ACD/labs Epizooties Chemins and Maule, provides information on common uses, chemical class, safety information, references, and links to other sources of information, is searchable by using a range of identifiers, and provides chemical classification codes physical property and toxicity data (eg. LD50 data for multiple species and routes) with links to original references. The Computational Toxicology Chemicals Dashboard can be searched by chemical identifiers (eg. AS number) product categories and assays/genes associated with high-throughput screening. It provides extensive information on chemistry, toxicity, and exposure data, including physical and chemical properties, environmental fate, usage, in vivo toxicity data, and results from a wide range of in vitro assays. The QSAR Toolbox has been developed to support read-across predictions. A significant number of databases have been donated to this project, hence, it represents a useful resource for human and environmental toxicity endpoint data as well as physicochemical property and metabolic data AMBIT also designed to support chemical safety assessment contains the REACH data from the European Chemicals Agency as well as the European Food Safety Authority (EFSA) Open Food Tox databases.

Data quality assessment

The scale of the resources outlined indicates the vast amount of data that are available from which in silico models can be built to predict properties of interest (eg. ADME or activity/toxicity). It is essential to ascertain the data to be used, as any model is only as good as the data on which it is built Equally important is the sequitur that no model can be more accurate than the data from which it is derived. In addition, biological data are inherently variable and this sets the upper limit for the accuracy of predictive models, as was recently exemplified in an investigation into the levels of uncertainty in models based on data from the US EPA's Toxicity Reference Database. Quality is a relative term the purpose for which the data are to be used dictates the minimum level of data reliability and relevance that would be considered acceptable this determines the suitability of the data for a given purpose (data adequacy). Definitions for the various terms that are associated with data

quality have been outlined previously but they are summarized below.

Validity of data can be defined as evaluating the method used to generate data relative to accepted guidelines or the extent to which the methods used find the truth as a result of the investigator measuring what they intended to measure.

Accuracy can be defined as the closeness of agreement between test method results and accepted reference values.

The reliability of data is linked to the reliability of the experiments carried out. For example, whether the results can be confirmed by comparison to standards and whether the methodology is repeatable.

-Relevance is the relationship between the test that is carried out and the effect that is of interest (i.e. the meaningfulness of the assay). For example, the highest quality data are required for the safety assessment of individual chemicals however, lower quality data may suffice for general screening or ranking of chemicals in product development.

Industry often works with specific types or subsets of chemicals therefore, models baih unang in-house data can be more relevant than those built using external data sets. In-house data sources are usually reliable and have the advantage of an audit trail for tracing or resolving issues.

Where necessary. One disadvantage in attempting to formally assess the quality of data is that the results can be highly subjective. The Klimisch criteria are the most widely used for classifying data quality. The application of these criteria enables data to be categorized as (i) reliable without restriction (ii) reliable with restriction (iii)unreliable or (iv) not assignable (ie. insufficient information exists on the data for a judgment of its quality to be made). The ToxrTool available from the European Commission's Joint Research Centre (JRC) is designed to help data users and modelers assign Klimisch quality scores to in vitro and in vivo toxicity data, by posing a series of questions relating to the methodological details. More recently, the Science in Risk Assessment and Policy (SciRAP) project has developed web-based tools for the evaluation and reporting of (eco)toxicity data to increase the structure and transparency of data reliability assessments.

The Criteria for Reporting and Evaluating Ecotoxicity Data (CRED) provide a means to characterize the quality of data for Ecotoxicological endpoints. The quality of the data used to build or evaluate a (Q)SAR model is a determinant of model quality and prediction reliability. Hence, appropriate precautions should be taken, such as checking for accuracy (e.g. avoidance of transcription errors in large compilations) reliability, and relevance by using primary sources where possible.

Structure-activity relationships and structural alerts

Stractive-activity relationships (SARs) refer to any definable relationship between a molecular feature of a chemical and activity. Simple rule-bus of clasification schemes cut-off criteria or generic rules-of-that are the simplest examples Lipinski's Rule of Fives which is designed to screen out drag candidates with potentially poor oral absorption, is probably the most well-known of these. Lipunki's Rule states that chemicals with a molecular weight above 500 Da, a logarithm of the octanol-water partition coefficient (log P) above 5, more than 5 hydrogen bond donors, or more than 10 hydrogen bond acceptors are associated with low oral absorption. Sample rules have also been developed for other properties of interest, for example, if the number of nitrogen pha oxygen atoms in a molecule is less than or equal to five it has the potential to penetrate the blood-brain barrier. While there are many exceptions to such generic rules many have been taken up widely for preliminary screening purposes notably in early drug development

There are numerous examples of software (freely available and commercial) that can generate simple physicochemical properties for chemicals and apply rules-of-thumb

Certain structural features may be associated with specific mechanics rules-of-thumb on giving rise to structural alerts.

Many structural alerts have been identified

REFERENCES

[1] Worth A. Computational modelling for the sustainable management of chemicals. Comput Toxicol 2020; 14: 100122.

[2] Mahony C, Ashton RS, Birk B, et al. New ideas for nonanimal approaches to predict repeated-dose systemic toxicity: report from an EPAA Blue Sky Workshop. Regul Toxicol

Pharmacol 2020; 114: 104668.

[3] OECD. Guidance document on the reporting of defined approaches to be used within integrated approaches to testing and assessment, Series on Testing and Assessment No.255. Paris: Organisation for Economic Cooperation and Development, 2016, 23 pp.

[4] WHO. Environmental health criteria 242: dermal exposure. Geneva: World Health Organisation, 2014, 503 pp.

[5] Cherrie JW, Fransman W, Heussen GAH, et al. Exposure models for REACH and occupational safety and health regulations. Int J Environ Res Public Health 2020; 17: 383.

[6] Cruz-Monteagudo M, Medina-Franco JL, Perez-Castillo Y, et al. Activity cliffs in drug discovery: Dr Jekyll or Mr Hyde? Drug Discov Today 2014; 19: 1069–1080.

[7] Cronin MTD and Yoon M. Computational methods to predict toxicity. In: Balls M, Combes R and Worth A (eds) The history of alternative test methods in toxicology. Cambridge, MA: Academic Press, Elsevier, 2019, pp. 287–300.

[8] Carnesecchi E, Toma C, Roncaglioni A, et al. Integrating QSAR models predicting acute contact toxicity and mode of action profiling in honey bees (A. mellifera): data curation using open source databases, performance testing and validation. Sci Total Environ 2020; 735: 139243.

[9] Piir G, Kahn I, Garc´ıa-Sosa AT, et al. Best practices for QSAR model reporting: physical and chemical properties, ecotoxicity, environmental fate, human health, and toxicokinetics endpoints. Environ Health Perspect 2018; 126: 126001.

[10] European Commission. Regulation (EC) no. 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council
Directive 76/769/EEC and Commission Directives 91/155/ EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. Off J Euro Union 2006; L349: 1–849.

[11] ECHA. The use of alternatives to testing on animals for the REACH Regulation. Fourth report under Article 117(3) of the