

In Silico Toxicological Approaches in Safety Estimation of Chemical Compounds with Stepwise Prioritization Model

Stanislava Georgieva

Department of Chemistry, Medical University - Varna, 55 Marin Drinov Str., 9002 Varna, Bulgaria

stgeorgieva@mu-varna.bg

Abstract: In the modern living conditions and in spite of the wide variety of existing pharmaceutical preparations, the search for new sources of biologically active compounds for the production of more effective and safer drugs remains relevant. The creation and management of alternative approaches to assess the safety of chemical compounds with potential activity, and combining information with experimental data and computational toxicological results are becoming more common and preferred models, carried out without the use of animals.

Keywords: in-silico, toxicology, chemical compounds.

Received 01 September 2023

Accepted 09 October 2023

Published 25 October 2023

1. Introduction

Drug discovery and development is a very lengthy and high-risk process, during which a complex set of activities are carried out, including analysis of information sources, synthesis of substances, conduct of pharmaceutical, pre-clinical and clinical studies, culminating in registration of the drug.

The main sources of medicinal substances are:

- compounds derived mainly from plant or microbiological sources;
- compounds obtained by chemical synthesis;
- compounds obtained by biotechnology and biochemical methods providing secondary metabolites and recombinant proteins by exploiting natural metabolic pathways and generating new ones.

Biological and bio-based products are becoming a large branch of pharmaceutical production with their own specificity.

Data indicate that 96% of drug candidates are dropped in the early stages [1] and the average cost of developing a new drug has reached up to USD 2.6 billion in recent years [2]. New molecules need to meet a number of medical and non-medical criteria: specificity (defined mechanism of action); efficacy (high affinity for a chosen target), selectivity for other possible biological targets and, above all, safety for patients. Intellectual property rights for the active molecule should be clearly defined and well protected; the compound should be available at a reasonable cost and a technically feasible process, easy to validate and control, consistently providing the active substance with high chemical purity and desirable physicochemical properties [3].

One of the main reasons for the high dropout rate in the early stages of research is safety. The safety assessment of drug candidates is most often carried out by pre-clinical tests – in vitro (on cell cultures) and in

vivo (animal models). The first type are high-throughput, based on interaction with target proteins or selected cell lines, which allow the affinity/activity of a large set of compounds to be tested very quickly (within hours), and the second type are low-throughput, working on tissues, organs or animals, in which pre-selected compounds are tested by more detailed procedures, over weeks or months.

2. In silico toxicological approach

In silico toxicological approach refers to the use of computational methods and models to predict the potential toxicity of chemical compounds. The following are the steps that can be taken to perform an in silico toxicological assessment [4]:

- *Data collection and preparation* – Collecting data on the chemical structure and properties of the compound, as well as any available toxicity data from previous experiments or studies. The data needs to be curated and standardized to ensure accuracy and consistency;
- *Selection of predictive models* – Choosing appropriate computational models for predicting toxicity based on the type of chemical and its intended use. There are various types of models that can be used, such as QSAR (Quantitative Structure-Activity Relationships), PBPK (Physiologically Based Pharmacokinetic) models, and read-across methods;
- *Model development and validation* – Developing and training the predictive models on the available data, and validating their accuracy and reliability. The models should be tested against independent datasets to ensure their robustness and consistency;
- *Prediction of toxicity* – Using the validated models to predict the potential toxicity of the chemical compound. The predicted toxicity can be compared to regulatory safety thresholds to determine if the compound is safe for its intended use;
- *Risk assessment* – Combining the predicted toxicity data with exposure data to assess the risk associated with the use of the compound. This involves



estimating the likelihood and severity of any adverse effects that may occur;

- *Reporting and interpretation of results* – The results of the *in silico* toxicological assessment should be reported in a clear and concise manner and the implications of the findings should be interpreted and communicated to stakeholders.

Overall, *in silico* toxicological approaches can provide valuable insights into the potential hazards and risks associated with chemical compounds, and can help guide decision-making around their safe use. However, it is important to note that these methods are not a replacement for traditional toxicology testing [5], and should be used in combination with other approaches for comprehensive safety evaluation.

2.1. *In-silico* research principles for toxicology

In silico toxicology studies follow the principles of the 3Rs (Replacement, Reduction and Refinement) proposed by Russell and Burch and support the creation and refinement of modern methods for assessing the safety of medicinal products.

Substitution refers to technologies or approaches that directly replace or avoid the use of animals in experiments where they would otherwise be used. Animal models are often expensive and time consuming and, depending on the research question, present scientific limitations such as a weak link to human biology. Over the last decade, advances in science and technology mean that there are now realistic options to replace the use of animals [6]. The replacement can be divided into two categories:

- Full replacement refers to methods that avoid the use of animals for research and testing purposes. It involves the use of human volunteers, tissues and cells, mathematical and computer models and established cell lines - often collectively referred to as animal-free technologies or NATs. In recent years, the term New Approach Methodologies (NAM) has been adopted by the bioscience sector specifically to describe non-animal technologies for use in assessing chemical or drug toxicity.

- Partial replacement involves the use of certain animals that, based on current knowledge, are not considered to be capable of experiencing suffering. This includes invertebrates such as *Drosophila*, nematode worms and social amoebae and immature forms of vertebrates. Partial replacement also involves the use of primary cells (and tissues) taken from animals killed solely for this purpose (i.e. not used in a scientific procedure that causes suffering).

Reduction refers to methods that minimise the number of animals used in an experiment or study, consistent with scientific objectives. It is essential for reduction that animal studies are appropriately designed and analysed to ensure robust and reproducible results.

Reduction also includes methods that allow the information collected on an animal in an experiment to be maximised to reduce the use of additional animals. Examples of this include the use of some imaging methods that allow longitudinal measurements of the

same animal (rather than, for example, culling cohorts of animals at specific points in time) or blood microsampling, where small volumes allow repeated sampling in the same animal. In these scenarios, it is important to ensure that the reduction in the number of animals used is balanced against any additional suffering that may be caused by their repeated use.

Enhancement refers to methods that minimize pain, suffering, distress or lasting harm that may be experienced by the animals being tested and that improve their welfare. Enhancement refers to all aspects of the use of animals, from their housing and husbandry to the scientific procedures carried out on them. Examples of refinement include providing animals with housing that allows the expression of species-specific behaviours, using appropriate anaesthesia and analgesia to minimize pain, and training animals to cooperate with procedures to minimize any distress.

Evidence suggests that pain and suffering can alter an animal's behaviour, physiology and immunology. Such changes can lead to variations in experimental results that impair both the reliability and repeatability of studies.

Compared to experimental approaches, computational methods have shown great advantages as they minimize testing time as well as the need for animal experiments and associated costs [7].

Innovative skills, methods, and processes, as well as novel ways to use established processes provided by artificial intelligence, can help research and regulatory entities overcome many of the identified challenges to data collection, processing, and analysis, and help participants identify and manage risks more effectively and efficiently in creating new and safe medicines. The advent of computational methods and processes complement conventional experiments and accelerate hypothesis generation and the research validation cycle.

Analyzing the numerical data of a huge number of new molecules with their characteristics enables their stratification. This is achieved by introducing technologies for handling big data collected from biomedical sources.

In silico toxicological investigations use approaches to help solve safety problems based on chemical-biological informatics. (Fig. 1) With *in silico* methods, hazards can be assessed at an early stage, eliminating the need for synthesis. This can be particularly important during drug development, where *in silico* prediction of a compound can guide the choice for lead compounds, not only in terms of efficacy but also safety. Testing can therefore be targeted to specific hazards before a compound is referred for preclinical safety testing. In preclinical evaluation, IST methods are actively used to ensure the safety of impurities, degradants, metabolites and excipients, which is particularly valuable in cases where the test material is difficult to synthesize [8]. Acute and chronic toxicity, including carcinogenicity, reproductive and developmental toxicity, mutagenicity, organ toxicity, irritation/corrosion, sensitisation and lethality are considered at this stage.

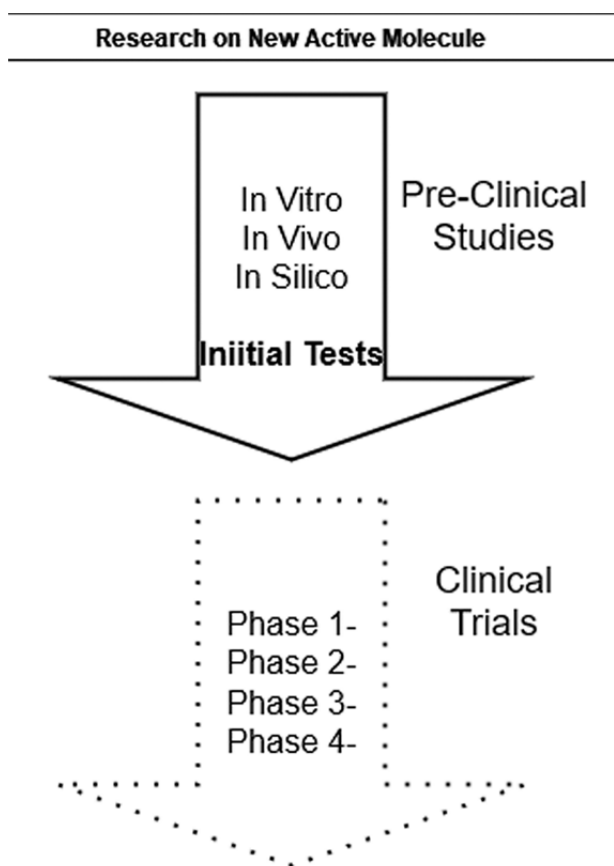


Fig.1. Position of the in silico research methods in pre-clinical studies.

2.2. In silico research rules and regulations

In silico toxicology studies follow consistent, transparent and well documented rules:

- 1) Analysis planning - identification of toxicological effects or mechanisms to be predicted; selection of in silico;
- 2) Database search and conduct appropriate individual software predictions;
- 3) Documentation of the in silico analysis and combined evaluation of the experimental data;
- 4) Standardized reporting of results and expert review against principles of Good Scientific Practice.

Toxicology-based computational approaches typically focus on toxicity database building, QSAR modeling, human rule-based methods, descriptor- and ligand-based methods, etc. There are even some data visualization tools available that are a hybrid containing toxicology prediction tools such as QSARs and systems biology and pharmacology pathway analysis [6].

2.3. Availability of in silico database

There are publicly available databases on toxicity, QSAR and effects on human health. With a great wealth of information there exist: AERS, ACToR, BDSM, CEBS, CEDI/ADI Database, CERES, Danish (Q)SAR Database, DevTox, Drugs FDA, DSSTox, EXTOXNET, TOXNET, and others. Toxicity databases are used to support the prediction of adverse effects of

pharmaceuticals and other xenobiotics for safety assessment, support lead selection and optimization, and ideally to add to safety or risk assessment high-quality test data from generated structures. The key concept for developing a drug and chemical toxicity database is the ability to use data of acceptable scientific quality from previous toxicity studies to build an electronic resource that can be searched, modeled, and used in "read-through" strategies with structurally related compounds(7) . Toxicity databases typically serve as predictive models or "mines" of data rich in authenticated structures, substance-induced toxicity experiments, or other similar scientifically based evidence.

2.4. In silico methodology

The main silico forecasting methodologies include the following:

- Statistical forecasting based on quantitative structure-activity relationships (QSAR). The method is based on the concept that molecules with similar structures potentially exhibit similar chemical and biological activities. The QSAR approach uses experimental data sets including the biological activity of chemical compounds; their chemical and physical characteristics represented as molecular descriptors and statistical methods to correlate these molecular descriptors with biological activity.

- Molecular descriptors are arithmetic values representing the physicochemical properties of compounds and can be categorized depending on the amount/type of information provided. The most common types of descriptors used in QSAR are constitutional, electronic, topological, and geometric descriptors and include molecular weight, total number of atoms, total number of carbon atoms, atomic network, total number of bonds, and van der Waals area (8). Over the years, the original QSAR analysis has been extended by various groups. Today, in the field of in silico toxicology, the term QSAR is very widely used to describe all forms of predictive modeling, including more complex (and less quantitative) machine learning models.

- Machine learning is a term encompassing supervised model building for classification and regression, unsupervised techniques such as clustering, and reinforcement learning techniques applied mostly to sequential decision tasks such as molecular design [9].

3. Cloud computing-enabled in silico research

Cloud computing has proven to be useful for in silico research, and its applicability for more highly specialized computing related to high performance computing (HPC) and similar applications is yet to be proven. There have been quantitative performance analyses performed on a set of benchmarks designed to represent a typical HPC workload on the Amazon EC2 cloud infrastructure. Their results show [10] a strong correlation between the time spent by an application in communication and its overall performance on EC2. It is also known that performance degrades as communications increases. In addition, the communication pattern of the application has a significant impact on its performance. Some applications with

significant global communication are reported to perform worse than those with less global communication. Thus, it appears that the amount of variability in environments such as EC2 cloud performance can provide significant gain. Variability is introduced by the shared nature of the virtualized environment, by the network, and by differences in the underlying non-virtualized hardware.

Virtual machines have been around and used for a very long time, the technology is widely used and is the basis of all modern cloud platforms. Furthermore, thanks to years of development and use, virtual machines have also proven themselves in a huge number of in-silico model trials. Container-based virtualization works by sharing the operating system kernel. Containers cannot completely replace virtual machines. There exists open source container platform that was explored to make containers easier to use for developers and system administrators. The renowned product comes with an easy to use container management interface and provides a robust ecosystem for sharing container applications in the form of templates, also called - images. Nowadays, more and more companies and educational institutions which perform large-scale in-silico research appear to be switching to the use of containers.

Current trends in industry pharmaceutical transformation require a transformation of both the education sector and the IT infrastructure used. There is a need for delivery models and the ability to deliver dynamic IT infrastructure using modern technologies where virtualization is at an unparalleled level compared to when it was invented. In future projects in the in-silico field, there would be the focus to investigate appropriate options for virtualization choices and to derive criteria that would be useful for training IT architects at the pharmaceutical companies and specialized universities, who are tasked with transforming or building infrastructure for the research needs of educational institutions.

Since then, virtualization and operating systems have been constantly evolving. Nowadays, the understanding is that virtualization is not used to run several simple tasks at the same time, but to run several software (such as operating systems) on the same physical machine without them interfering with the main operating system of the machine. There are many options for virtualization. Each user, whether an individual or an organization, has completely different needs. Nowadays, there is a wide range of solutions that fit the different situations where isolation is more or less needed.

Until recently, the typical method of virtualization, among individual users and organizations, was the use of virtual machines provided by hypervisors developed by different companies - Oracle (VirtualBox), VMWare (ESXi) or Microsoft (Hyper-V). The main advantage of full virtualization is that isolation is at the highest level from the host machine (the machine supporting the virtualization), providing the ability to have a wider choice of supported operating systems and a clear separation between the running host operating system and the virtualized one, providing a safer environment. Nevertheless, virtual machines have a major disadvantage - the amount of resources required for virtualization is

higher, as it is a completely isolated system that does not share its resources with the host OS kernel, whereas containers share the host OS kernel.

Given the performance of servers today, it is almost impossible to use their full potential with a single application. This is where the ability to install more than one application on each server comes in. However, this technology has certain risks. What happens when one application needs specific libraries or an entire environment with a certain version, and another needs something different? Then virtualization is the tool that helps the most. Using virtualization, we can run many applications on one machine without them being aware of each other.

Virtualization does not solve all problems. One of the main problems with virtualization servers is the difficulty of scaling, as there is no easy way to automatically react when resource needs or application requirements change. A simple example: let's partition a website and create a virtual environment for it that we deploy on two servers (due to replication in case one machine fails). Along with this application there can be many others running in their virtual spaces. When traffic rises to unexpected levels, the application can quickly demand more resources than the virtual machine can provide, and since there are more virtual machines on the server, the resource allocation cannot be increased. Therefore, the entire virtual machine must be moved to another physical machine that has more resources. Even after moving, this situation may occur again a few hours later when this application will not need many resources and another will be busier.

With the rapid development of technology, the main focus of IT companies is on virtualization, load balancing, scaling the deployment process, managing dependencies, ensuring stability across platforms, cyber security, etc. This can now be achieved through containerization. Containerization offers an alternative virtualization method where a single operating system (OS) on a host can run many different applications from the cloud. Compared to virtual machines, containers provide the ability to virtualize the OS. Containers offer a logical encapsulation mechanism where applications can be abstracted from the environment in which they actually run. This separation allows containers to be deployed easily and consistently, regardless of the target environment.

Software containers offer a different type of virtualization. The advantage of containers is that they offer a reduced level of isolation (compared to virtual machines), at an affordable cost, as they use and share resources with the host OS kernel already running. Docker, as an open source platform, provides abstraction layers to help the user leverage this type of virtualization. Because of its design, it offers a convenient way for users to create new containers and make applications "portable", which is to signify that a given container can be run on any other operating system on which Docker is installed. This eliminates the need to deal with dependencies needed for the software to run when moving to another machine.

The Docker platform has evolved significantly in recent years. The fact that all products are open source

makes the community play a huge role in choosing what functionality to develop with priority in the next stage. The user base continues to grow and many large companies are publicly announcing that they are actively using Docker as part of their own infrastructure. However, this rapid evolution makes Docker a product that cannot yet be considered stable and reliable (compared to traditional virtual machines) due to the rapid pace of change.

4. Usage of datasets for silico toxicological research

In silico toxicology studies use computational methods to predict the toxicity of chemicals and other substances. The accuracy of these predictions depends

largely on the quality of the data used to train the models. This requires defining the scope of the study, which includes identifying the specific research questions that the in silico toxicology study aims to answer, as well as the types of chemicals or substances to be analyzed and the toxicological endpoints of interest.

Identifying the specific research questions that the in silico toxicological study aims to address is a critical step in the study design process and depends on the objectives of the study and the toxicological characteristics of interest. These may be : determination of potential carcinogenicity, mutagenicity, genotoxicity or endocrine disruption as a consequence of candidate drug or chemical use.

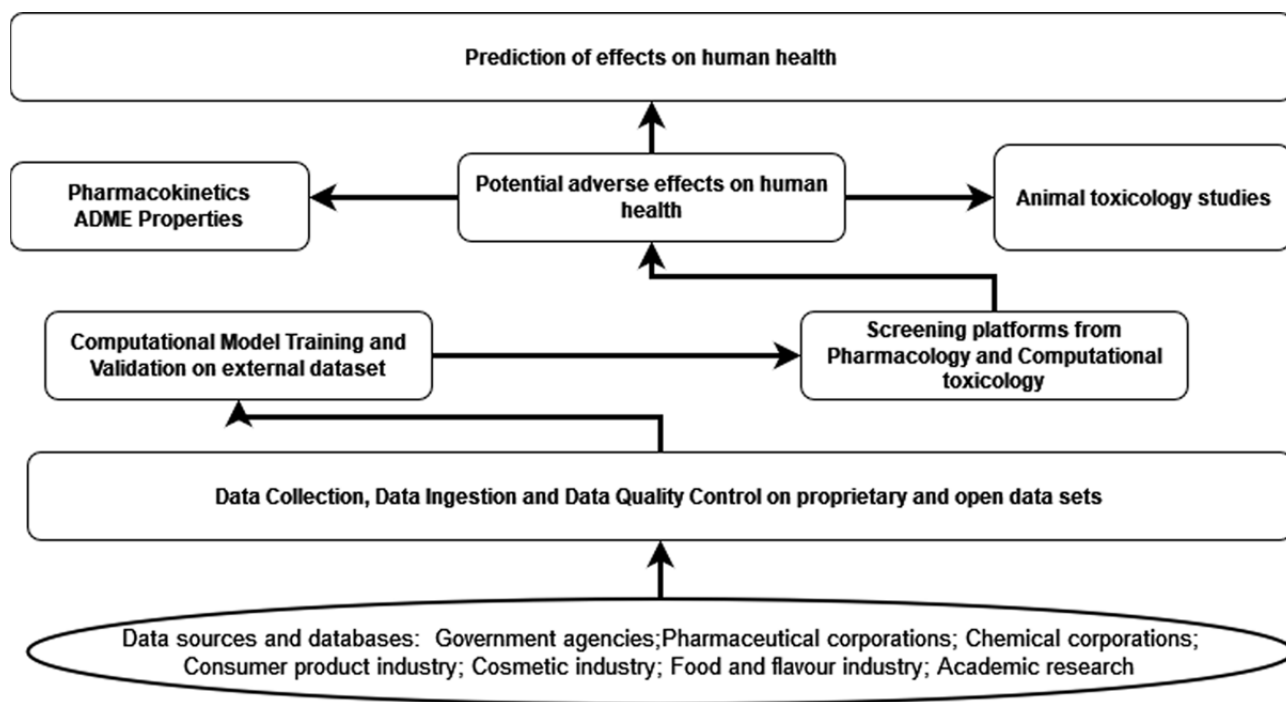


Fig.2. Processes and data workflow of data within in silico toxicology research.

Following the scoping of the toxicology study, a data collection and preparation stage for in silico toxicological studies follows. There are various sources of information, such as public databases:

- Tox21 [17]: A database maintained by the US Environmental Protection Agency (EPA) that contains data on the biological activity of chemicals.
- PubChem [16]: A database maintained by the National Institutes of Health (NIH) that contains information on the chemical properties and biological activity of small molecules.
- ChEMBL [18]: A database maintained by the European Bioinformatics Institute (EBI) that contains information on bioactive molecules and their targets.

Proprietary databases are typically held by companies that maintain their own databases of chemical toxicity data that they can use to support their own research or product development efforts.

Toxicity data collected and analyzed from published animal studies, in vitro and human studies, and

quantitative structure-activity relationship models or the expertise of toxicologists.

Quality control of the collected toxicological databases is an important step to ensure that the toxicological databases used in in silico studies are of high quality and suitable for generating accurate predictive models.

The skilful combination of all these sources of information, as well as their quality and completeness, serves to build successful predictive models and identify the toxicological characteristics of interest.

5. Approach for experiment conducting for in-silico toxicology research

The characteristics or variables that are most appropriate for in silico toxicology studies may vary depending on the predicted endpoint of the test and/or the specific data set used. Common characteristics to include are [11]:

Molecular Descriptors - These are numerical values that are used to describe various aspects of the molecular

structure of a chemical. In *in silico* toxicity prediction, molecular descriptors are used to develop quantitative structure-activity relationship (QSAR) models that can predict the toxicity of chemicals based on their molecular structures. The molecular descriptors can be calculated using various of computational methods including quantum mechanics, molecular mechanics, and molecular dynamics simulations. Some common types of molecular descriptors include:

- Topological descriptors- These are based on a graphical representation of a molecule and include information about the number of atoms, bonds, and rings in the molecule.
- Electron descriptors- Determined by the electronic properties of the molecule, such as its ionization potential or electron affinity.
- Geometric descriptors- Depend on the geometry of the molecule, such as its shape or symmetry.
- Constitutional descriptors- These descriptors are based on the composition of a molecule, such as the number and type of atoms and functional groups.
- Hydrophobicity descriptors- depend on the ability of the molecule to interact with water and include information on its partition coefficient and hydrophobic surface area.

1. Bioactivity data - Used to develop quantitative structure-activity relationship (QSAR) models that can predict the toxicity of chemicals/drug candidates based on their bioactivity profiles. Commonly used bioactivity profiles in toxicological studies are :

- Receptor binding affinity - Describes the ability of a chemical to bind to specific receptors in the body, such as estrogen receptors or androgen receptors, and cause an endocrine disorder.
- Enzyme inhibition - Defines the ability of a chemical to inhibit specific enzymes in the body, such as cytochrome P450 enzymes or acetylcholinesterase, and to exert adverse effects on metabolic processes.
- Ion channel modulation- There are cases where a chemical can modulate the activity of ion channels in the body, such as voltage-dependent sodium or calcium channels. This is used to predict the potential of a chemical to cause cardiac toxicity or neurotoxicity.
- Transport activity- Knowledge of a chemical's potential interactions with specific transporters in the body, such as P-glycoprotein or organic anion transporters, allows prediction of drug-drug interactions or the ability to accumulate in specific tissues.

2. Physicochemical properties are commonly used as descriptors in *in silico* toxicological studies to develop quantitative structure-activity relationship (QSAR) models that can predict the toxicity of chemicals. Some commonly used physicochemical properties include:

- Molecular weight - a measure of the mass of a molecule and can be used to predict properties of a chemical, such as absorption and distribution.
- LogP - indicates the hydrophobicity of the chemical and can be used to predict the ability of the chemical to cross biological membranes.
- Polar surface area - property tells us the surface area of a molecule that is polar or can form hydrogen

bonds, and can be used to predict a chemical's ability to interact with biological systems.

- Constant of acid dissociation (pKa)-A measure of the acidity or basicity of a chemical and can be used to predict the ionization state and solubility of a chemical under various pH conditions.

Water solubility- The ability of a chemical to dissolve in water and can be used to predict the toxicokinetic properties of a chemical and the potential for accumulation in the environment, etc.

Other important variables used in toxicological prediction are structural signals referring to substructures or functional groups known to be associated with specific toxicological effects (For example, the presence of a nitro group or epoxide ring in a structure serves as a signal to predict mutagenicity or genotoxicity). Exposure data, on the other hand, refers to information on how a chemical is likely to be encountered or absorbed by humans or other organisms. This may include information on routes of exposure, such as inhalation or ingestion, as well as information on the frequency and duration of exposure.

The choice of an appropriate method of predicting toxicity is a critical step in the development of *in silico* toxicological studies and is determined primarily by the factors considered so far, such as - structure and type of test substance, exposure and the available database for predicting acute toxicity, chronic toxicity, genotoxicity or other specific toxic effects. For this purpose, the training of predictive models is necessary, which is done by converting the collated information into a format that is convenient for analysis. The features or discriminators [12] to be used for model training are also identified. This involves selecting a set of physicochemical, biological or structural properties, which were discussed in above. After model construction, it is essential to confirm its performance or validate it. Model validation can be performed using a set of statistical metrics, such as accuracy, sensitivity, specificity, and area under the curve (AUC) [13].

Key steps for predicting the potential toxicity of new compounds include:

1) Inputting the chemical structure of the compound into the predictive model using computer fractions or using identification numbers.

2) Generation and recognition of molecular descriptors that can be used as input features for the predictive model. These descriptors may include physicochemical properties, structural features, and biological activities.

3) Apply the model and generate a prediction of the toxicity of the compound. The output of the model may be a binary classification (e.g., toxic vs. nontoxic) or a quantitative prediction of toxicity (e.g., LD50 or IC50). A concurrent risk assessment is performed to determine the likelihood and potential consequences of exposure to the chemical based on the toxicity prediction generated by the model.

4) Evaluate the model and compare the predicted toxicity with experimental data, if available, or assess the uncertainty or confidence in the prediction. The results of the prediction can be used to prioritize compounds for

further testing, to identify potential mechanisms of toxicity, or to support regulatory decision making.

6. Conclusion

The main features of in silico methods that make them preferable to in vitro and especially in vivo studies are:

- higher throughput;
- lower cost;
- less time consuming;
- continuous optimization possible;
- higher reproducibility if the same models are used;

- low compound synthesis requirements;
- have the potential to reduce the use of animals

Disadvantages include:

- quality and transparency of experimental data from the training set;
- transparency of the programme;
- sometimes confusing descriptors;
- area of applicability sometimes not clear;
- specific features of metabolism not taken into account;
- prediction of carcinogenicity is under refinement.

To date, positive experiences have been reported with in silico toxicology studies mainly for mutagenicity, sensitization which are areas with relatively well known mechanisms. Hepatotoxicity, neurotoxicity cannot yet be accurately predicted with in silico methods. The perspective here lies in the separation of complex endpoints into different steps or pathways with the common problem of how to validate and unify them to make a single prediction.

Acknowledgements

This paper is supported by grant Nr. 30-363#3/03.06.2022 “Research laboratory establishment for artificial intelligence and cloud technologies” by the University of Telecommunications and Post, Sofia, Bulgaria.

References

- [1] A.M.C. Sean Ekins, Alexander L. Perryman, Joel S. Freundlich, Alexandru Korotcov, Valery Tkachenko, Accessible machine learning approaches for toxicology, Ekins (Ed.), Comput. Toxicol (2018).
- [2] Baskin II, D. Winkler, I.V. Tetko, A renaissance of neural networks in drug discovery, Expert Opin Drug Discovery, 11 (2016), pp. 785-795.
- [3] G. Klambauer, T. Unterthiner, A. Mayr, S. Hochreiter, DeepTox: toxicity prediction using deep learnin, Toxicol. Lett., 280 (2017).
- [4] K.D. Wu, G.W. Wei, Quantitative toxicity prediction using topology based multitask deep neural networks, J. Chem. Inf. Model., 58 (2018), pp. 520-531.
- [5] Pearl G.M., Livingston-Carr S., Durham S.K., Integration of computational analysis as a sentinel tool in toxicological assessments, Curr. Top. Med. Chem., 1 (2001), pp. 247-255.
- [6] Lahl U., Gundert-Remy U., The use of (Q)SAR methods in the context of reach, Toxicol. Mech. Methods, 18 (2008), pp. 149-158.
- [7] Tanner R (2018) The 3Rs: What are Medical Scientists Doing about Animal Testing?. Front. Young Minds. 6:44. doi: 10.3389/frym.2018.00044.
- [8] J.C. Graham, M. Rodas, J. Hillegass, G. Schulze, The performance, reliability and potential application of in silico models for predicting the acute oral toxicity of pharmaceutical compounds, Regul. Toxicol. Pharmacol., 119 (2021), Article 104816, 10.1016/j.yrtph.2020.104816.
- [9] J. Bercu, M.J. Masuda-Herrera, A. Trejo-Martin, C. Hasselgren, J. Lord, J. Graham, M. Schmitz, L. Milchak, C. Owens, S.H. Lal, R.M. Robinson, S. Whalley, P. Bellion, A. Vuorinen, K. Gromek, W.A. Hawkins, I. van de Gevel, K. Vriens, R. Kemper, R. Naven, P. Ferrer, G.J. Myatt, A cross-industry collaboration to assess if acute oral toxicity (Q)SAR models are fit-for-purpose for GHS classification and labelling, Regul. Toxicol. Pharmacol., 120 (2021), Article 104843, 10.1016/j.yrtph.2020.104843.
- [10] SOKOLOV, Strahil; VLAEV, Stefan; ILIEV, Teodor B. Technique for Improvement of Backup and Restore Strategy Based on Blockchain. In: 2022 International Conference on Communications, Information, Electronic and Energy Systems (CIEES). IEEE, 2022. p. 1-6.
- [11] J. Harvey, A. Fleetwood, R. Ogilvie, A. Teasdale, P. Wilcox, S. Spanhaak, Management of organic impurities in small molecule medicinal products: deriving safe limits for use in early development, Regul. Toxicol. Pharmacol., 84 (2017), pp. 116-123, 10.1016/j.yrtph.2016.12.011.
- [12] K. Roy, S. Kar, R.N. Das, Chemical Information and Descriptors, Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment, Academic Press, Boston (2015), pp. 47-80.
- [13] J.B.O. Mitchell, Machine learning methods in chemoinformatics, Wiley Interdiscip. Rev.: Comput. Mol. Sci., 4 (2014).
- [14] Y. Cong, B.-k. Li, X.-g. Yang, Y. Xue, Y.-z. Chen, Y. Zeng, Quantitative structure–activity relationship study of influenza virus neuraminidase A/PR/8/34 (H1N1) inhibitors by genetic algorithm feature selection and support vector regression, Chemom. Intell. Lab. Syst., 127 (2013).
- [15] Raies, A.B. and Bajic, V.B. (2016), In silico toxicology: computational methods for the prediction of chemical toxicity. WIREs Comput Mol Sci, 6: 147-172. <https://doi.org/10.1002/wcms.1240>
- [16] Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., Li, Q., Shoemaker, B. A., Thiessen, P. A., Yu, B., Zaslavsky, L., Zhang, J., & Bolton, E. E. (2023). PubChem 2023 update. Nucleic Acids Res., 51(D1), D1373–D1380. <https://doi.org/10.1093/nar/gkac956>
- [17] Ann M. Richard, Ruili Huang, Suramya Waidyanatha, Paul Shinn, Bradley J. Collins, Inthirany Thillainadarajah, Christopher M. Grulke, Antony J. Williams, Ryan R. Lougee, Richard S. Judson, Keith A. Houck, Mahmoud Shobair, Chihae Yang, James F. Rathman, Adam Yasgar, Suzanne C. Fitzpatrick, Anton Simeonov, Russell S. Thomas, Kevin M. Crofton, Richard S. Paules, John R. Bucher, Christopher P. Austin, Robert J. Kavlock, and Raymond R. Tice, Chemical Research in Toxicology 2021 34 (2), 189-216, DOI: 10.1021/acs.chemrestox.0c00264
- [18] Mendez D, Gaulton A, Bento AP, Chambers J, De Veij M, Félix E, Magariños MP, Mosquera JF, Mutowo P, Nowotka M, Gordillo-Marañón M, Hunter F, Junco L, Mugumbate G, Rodríguez-Lopez M, Atkinson F, Bosc N, Radoux CJ, Segura-Cabrera A, Hersey A, Leach AR. Nucleic Acids Res. 2019; 47(D1):D930-D940. doi: 10.1093/nar/gky1075