



# Artificial intelligence and machine learning disciplines with the potential to improve the nanotoxicology and nanomedicine fields: a comprehensive review

Ajay Vikram Singh<sup>1</sup> · Mansi Varma<sup>2</sup> · Peter Laux<sup>1</sup> · Sunil Choudhary<sup>3</sup> · Ashok Kumar Datusalia<sup>2</sup> · Neha Gupta<sup>4</sup> · Andreas Luch<sup>1</sup> · Anusha Gandhi<sup>5</sup> · Pranav Kulkarni<sup>6</sup> · Banashree Nath<sup>7</sup>

Received: 15 January 2023 / Accepted: 20 February 2023 / Published online: 7 March 2023  
© The Author(s) 2023

## Abstract

The use of nanomaterials in medicine depends largely on nanotoxicological evaluation in order to ensure safe application on living organisms. Artificial intelligence (AI) and machine learning (ML) can be used to analyze and interpret large amounts of data in the field of toxicology, such as data from toxicological databases and high-content image-based screening data. Physiologically based pharmacokinetic (PBPK) models and nano-quantitative structure–activity relationship (QSAR) models can be used to predict the behavior and toxic effects of nanomaterials, respectively. PBPK and Nano-QSAR are prominent ML tools for harmful event analysis that is used to understand the mechanisms by which chemical compounds can cause toxic effects, while toxicogenomics is the study of the genetic basis of toxic responses in living organisms. Despite the potential of these methods, there are still many challenges and uncertainties that need to be addressed in the field. In this review, we provide an overview of artificial intelligence (AI) and machine learning (ML) techniques in nanomedicine and nanotoxicology to better understand the potential toxic effects of these materials at the nanoscale.

**Keywords** Artificial Intelligence (AI) · Nanomedicine · Physiologically based pharmacokinetic (PBPK) models · Nanotoxicology · Adverse outcome pathway (AOP) analysis · Machine Learning (ML)

---

Ajay Vikram Singh and Mansi Varma contributed equally.

✉ Ajay Vikram Singh  
Ajay-Vikram.Singh@bfr.bund.de

- <sup>1</sup> Department of Chemical and Product Safety, German Federal Institute for Risk Assessment (BfR), Max-Dohrn-Straße 8-10, 10589 Berlin, Germany
- <sup>2</sup> Department of Regulatory Toxicology, National Institute of Pharmaceutical Education and Research (NIPER-Raebareli), Lucknow 229001, India
- <sup>3</sup> Department of Radiotherapy and Radiation Medicine, Institute of Medical Sciences, Banaras Hindu University, Varanasi 221005, India
- <sup>4</sup> Department of Radiation Oncology, Apex Hospital, Varanasi 221005, India
- <sup>5</sup> Elisabeth-Selbert-Gymnasium, Tübinger Str. 71, 70794 Filderstadt, Germany
- <sup>6</sup> Seeta Nursing Home, Shivaji Nagar, Nashik, Maharashtra 422002, India
- <sup>7</sup> Department of Obstetrics and Gynaecology, All India Institute of Medical Sciences, Raebareli, Uttar Pradesh 229405, India

## Introduction

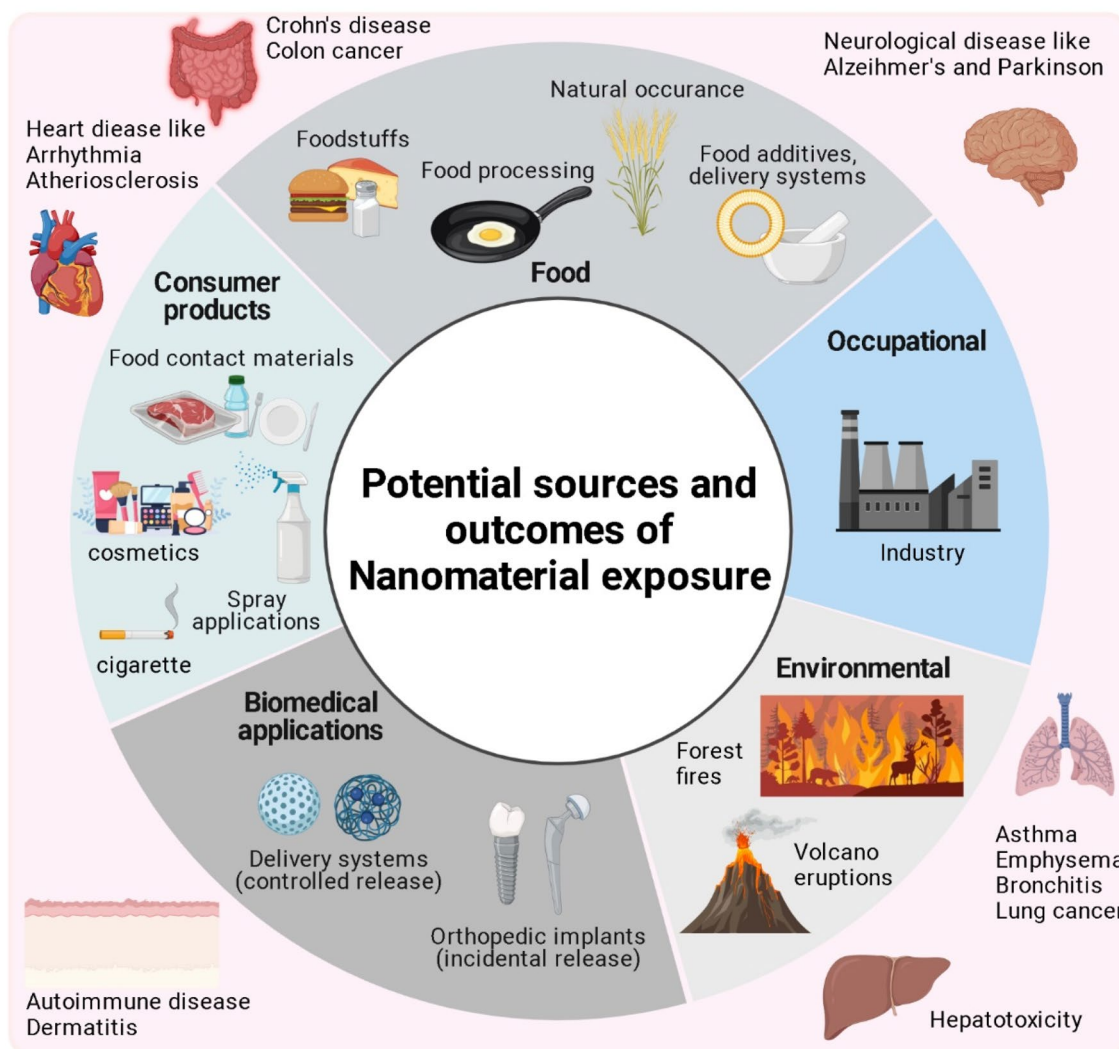
The nanomaterials are the most important advancement in science and technology worldwide. This invisible small size particles between 1 and 100 nm range have unique physical, chemical and biological properties which has applications in a wide range of fields (Bayda et al. 2019). According to a report by MarketsandMarkets, the global nanomaterials market is expected to reach \$75.64 billion by 2025, growing at a compound annual growth rate of 13.2% from 2020 to 2025. Nanomaterials is widely used in everyday life such as sunscreen, cosmetics, food packaging, water filtration, medicine and energy production. Nanotechnology and nanomedicines have given world a wide range of benefits and will continue to do so but it has become highly necessary to address its undesirable effects. Nowadays, the world is facing huge number of diseases originating from the daily exposure of harmful chemicals or materials whose behavior is not known hiding behind their application and benefits (Domingues et al. 2022). We are exposed to nanomaterials through industries, food

additives, processed food, cigarette, cosmetics, packaging materials, forest fires, controlled release medications, propellants, paints, etc. These have the potential to produce risk to human health and cause diseases like Parkinson, Alzheimer, asthma, cancer, emphysema, bronchitis, arrhythmia, dermatitis, vasculitis, urticaria, crohn's disease, hypertension, thrombosis, podoconiosis and many more (Fig. 1) (Ahamed 2014; Asati et al. 2021). Hence, for a safer and healthier future, there is a need to address this issue and do efficient nanotoxicological testing. Since the market on nanoparticles is growing rapidly, there are huge number of nanoparticles created and are going to increase which is little difficult to regulate for its safety through in vivo and in vitro tests. With the high speed and volume data growing, it is becoming difficult to assess chemicals using traditional methods which is even more

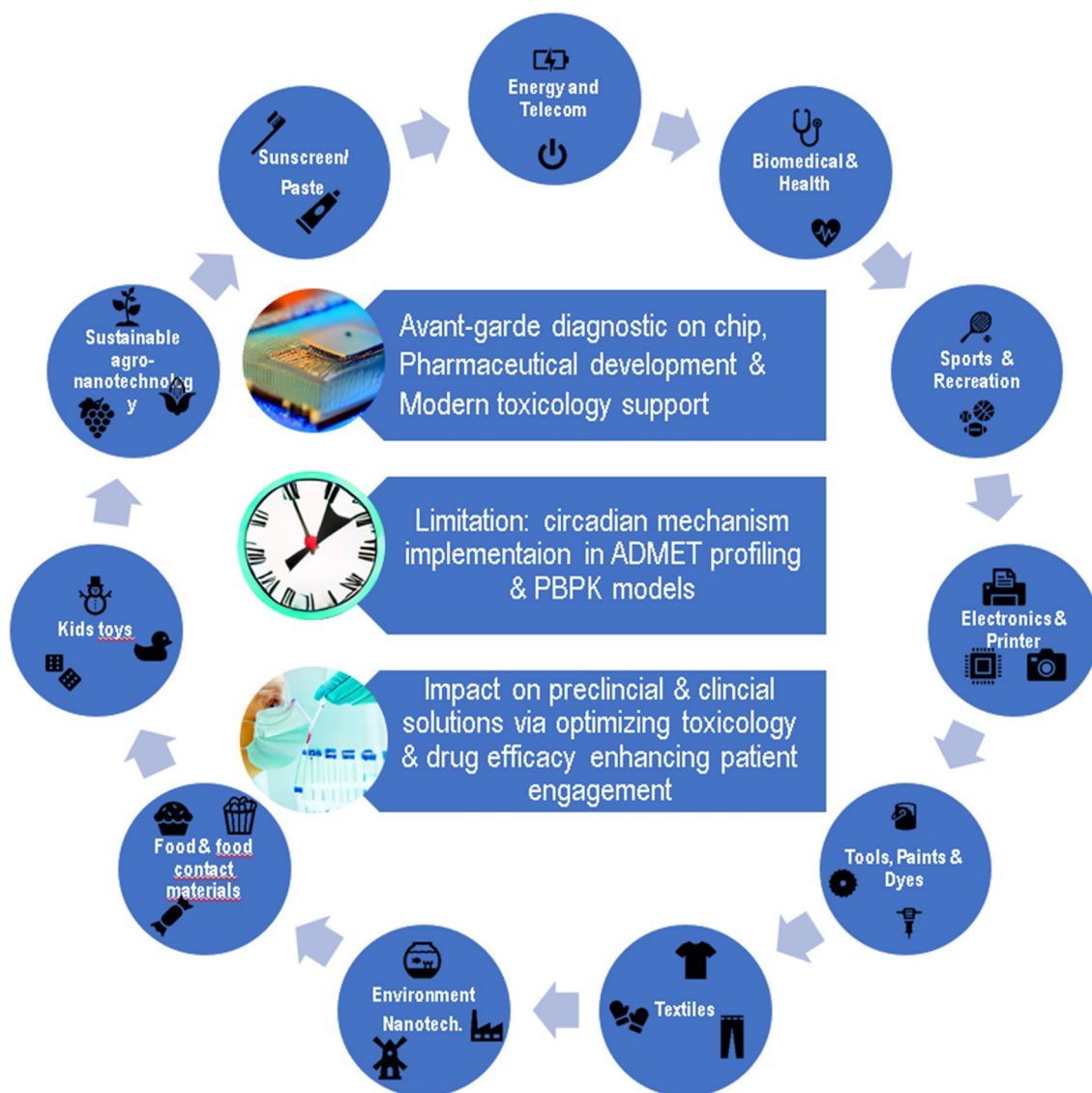
challenging to assess with various number of chemical toxicological endpoints (Singh et al. 2020a). Therefore, to meet the need to assess/predict the risk of such particles, we should opt for computational modeling methods that will save time and resources and help build community with better health.

### AI/ML modeling approaches for nanotoxicology related to systems biology and bioinformatics

The science of artificial intelligence is evolving in revolutionizing way and has a significant impact on our lives. It has the potential to contribute significantly in number of fields including healthcare, finance, transportation, and



**Fig. 1** The potential sources (depicted in inner circle) and the health outcomes (depicted in outer square) of nanomaterial exposure (Ahamed 2014). This figure was made using Biorender.com



**Fig. 2** AI-ML application in wide sectors and their intrinsic worth in shaping human life

manufacturing as well as resolving challenges like environmental protection, disaster response and social issues (Fig. 2). The AI is advancing and it has provided significant approaches to improve the process of drug discovery and development (Fig. 3). Research toward the field of systems biology and bioinformatics is focused toward the assessment of the adverse effects of the chemicals. Using bioinformatic tools and modeling, we can predict or can explain the adverse effects associated with any chemical.

AI enables to develop text mining and data mining approaches as a complement to more network biology and system biology approaches to understand mechanisms of diseases and the mechanisms of chemical actions (Kumar and Saha 2022). In this section, we explain briefly about shortcomings of chemical with respect to (w.r.t) consumer

safety, currently in why we need approaches from the modeling field to understand that new chemicals entering in market can produce an adverse effect. More than 500 years ago as said by Paracelsus, “What is there that is not poison? All things are poison and nothing is without poison. Solely the dose determines that a thing is not a poison” who is in fact considered the father of toxicology and this is really key today and also current knowledge today that the importance of the dose of the compound to produce a toxic effect (Grandjean 2016).

Later, it was realized that it is not only the dose that is important but also the time dynamics and that is why it's important that we are able to apply different modeling approaches for understanding the mechanisms of chemical toxicity. As a background why researches are interested in

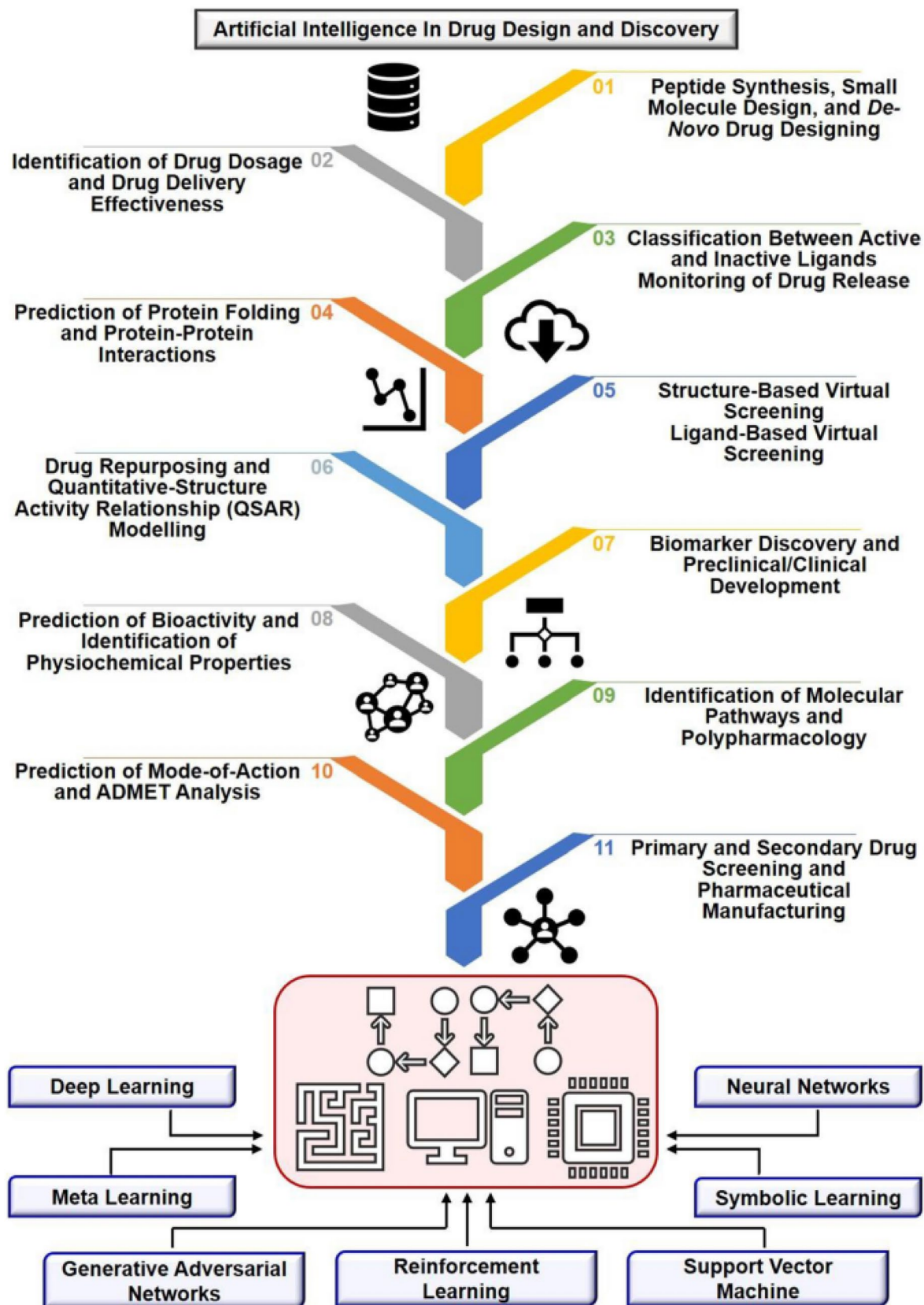


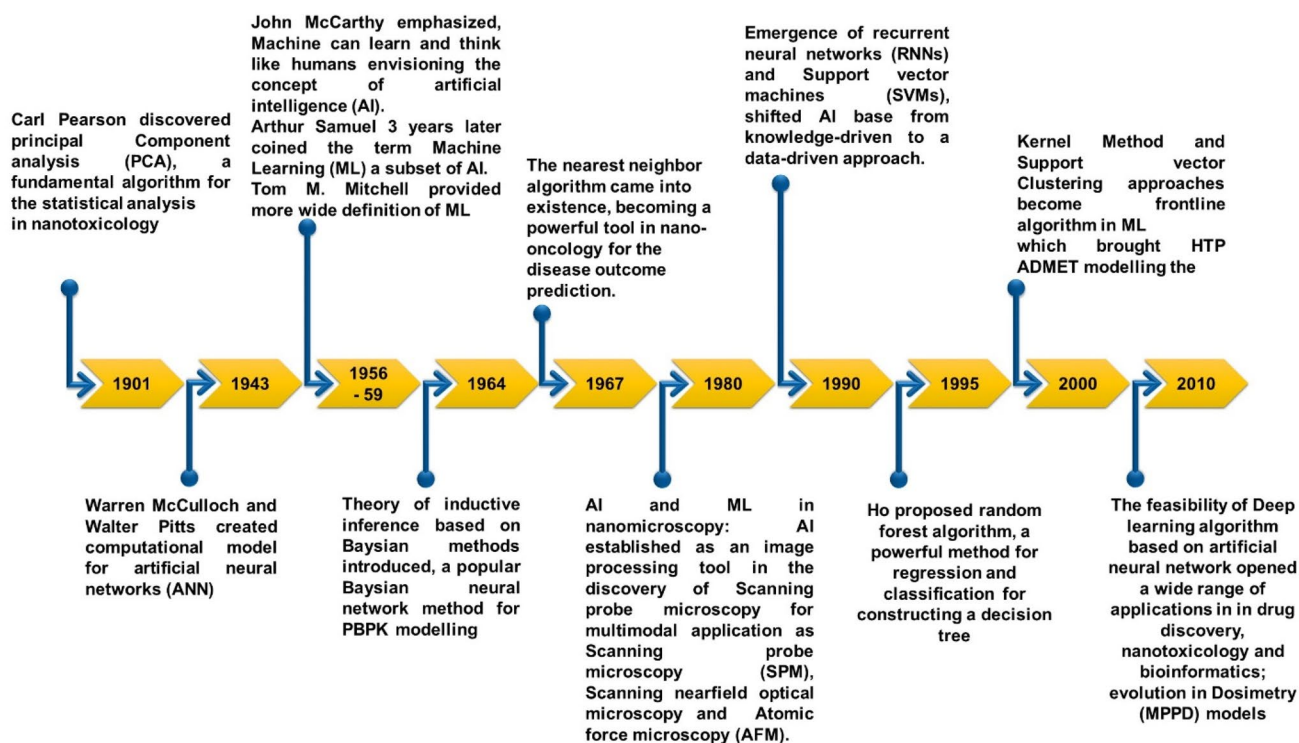
Fig. 3 Artificial intelligence and machine learning in drug discovery and development (Gupta et al. 2021)

chemical safety, as it is one of the major causes of attrition of new chemical during the whole process of chemical discovery and also after the chemical is marketed in context with regulatory affairs (Fuelle and Lanctin 2022). That is why the adverse reactions of chemicals are a very important concern in the clinical setting during the new area of application in consumer market and how can we predict these adverse chemical reactions? In general, the preclinical animal testing are failing to predict the human adverse chemical reactions, because they are only able to explain or predict this in 30% of the cases (Singh et al. 2021c). The main reason of that is because we still do not understand what are the mechanisms, why and how a chemical produces a chemical toxic and sub-toxic effect (e.g. genotoxicity, mutagenicity, etc.) and in fact for many chemicals which are there in the market, we do not have a detailed understanding of the mode of toxicological action and for asymptomatic adverse effect that is even worst (Singh et al. 2021b). Henceforth, it is important to realize that we are facing in our quest to understand the mechanisms of chemical toxicological mode of action and in particular chemical toxicity. We are addressing a multi-scale problem, because chemicals are in continuity exposed at organism level via different routes ca. dermal, ocular, inhalation and oral route (Chandrasekar et al. 2022).

Subsequently, chemicals exert their action at the molecular level via chemical targets such as proteins, peptides, DNA, RNA or other molecules at the cell. Later, they translate the effect to tissues/organ level via specific cellular mechanisms (Singh et al. 2020b). Therefore, we need to cover all these scales at the functional level as well as at the temporal level to study the mode of action of tracks and also to understand the adverse effects.

## Quantitative systems toxicology

Looking at history of toxicology, the QST started the testing of the toxic effect of chemicals and also chemical compounds in general in early in the last century when many people died from nephrotoxicity i.e. toxicity in the kidney due to the use of antibiotic (Petejova et al. 2019) (Fig. 4). Therefore, this incident prompted the testing of toxicity of compounds by using animal models and this is currently current practice in pharmaceutical industry before testing these compounds in humans and also from several legislation to regulate the toxicity of the compounds. A second very important event in the 1950s and 1960s was that of chemical Thalidomide which was used to relieve morning



**Fig. 4** Development of nanomedicine & toxicology over time using AI and ML. Timeline for the evolution of AI, which includes statistical methods, as well as nanoparticles (NPs), commencing with the first synthesis and quantum effects discovered by Faraday in 1853.

When AI was used for tasks like the identification of NP properties or interaction partners, the grouping of NPs based on their qualities or harmful effects, and the prediction of NP toxicity in 2010, both time-frames merged

sickness and other types of symptoms in pregnant women produced teratogenic effects and more than 10,000 cases were reported (Kim and Scialli 2011). It is important to note that these toxic effects for Thalidomide could not be predicted by the animal studies that were conducted contemporarily in rat (Swaters et al. 2022). All these events highlighting it for a more systemic approach to do toxicology and the other facts that are have all these developments are what we already seen but is also current today that there is a poor translation from animal experiments to the human in vivo scenario. From that, it is obvious that not all the animal models are good models or good predictors to what clinicians will observe then in patients. That is why we need to address this and computational approaches are playing important role in modern toxicology (Hemmerich and Ecker 2020). Also, the cost is determining factor since there is a lot of resource investment to approve and develop new chemicals risk assessment and also another important issue is that currently a lot of animals are used for toxicity testing and there is a movement toward decreasing the use of animal testing in different types or different kinds of toxicity (3Rs principle) (Granath et al. 2014). This is the reason in the last decade, the field of quantitative systems toxicology emerged and this is the definition that we have that “the goal is to provide a quantitative understanding”. That is why we mentioned before the importance of those in time in the response of the toxic effect of a chemical and in the organisms for going from the molecular to the phenotypic observation. It is done by integration of computational approaches and different experimental methods. This is a broad approach that can be applied to a different kind of chemicals in particular can be applied to environmental toxicity testing better will focus today to their application for in the field of chemical risk assessment (Pérez Santín et al. 2021). In this area, it has some particular implications. In this section, we would like to summarize a little bit what are the mechanisms of or the key events of chemical actions and in how they produce adverse effects.

### Essential role of PBPK models and ADMET profilers in health hazard predictions

The chemical can have a desired effect by acting on various targets. It is important to note that some chemicals need to be metabolized into active compounds before they can act on the targets and produce the desired effect (Yu et al. 2018). There is also the potential for the chemical to produce other harmful effects. This can be directly caused by the action of the chemical or its metabolites on targets. Additionally, detoxification processes in the liver, such as the ADMET profile, convert the chemical metabolites into byproducts, which can be extracted from the body.

These products can be chemically reactive and lead to DNA adducts modifications, mutations, and disruption of enzyme reactions. Many current QSAR models, molecular docking tools and ADMET properties prediction tool in silico are available to weigh the toxicity of chemicals (Daoud et al. 2021). Table 1 shows the comprehensive list of ADMET profiling tools, the parameters one can predict with their link adopted with permission from references (Shin et al. 2017). ADMET profilers and PBPK models are invaluable resources for connecting chemical toxicity and exposure data. They are essential for combining animal, in vitro, and computer-based experiments to aid in chemical assessments. Utilizing QSAR metabolic simulators, you can investigate if there are any known or simulated metabolites or hydrolysis products of the target chemical(s) (Yordanova et al. 2019).

In this section, we will discuss the adverse effects of chemical metabolites, which include reactive products that can form adducts with other molecules. An example of this is conjugation with glutathione, which can lead to direct damage or deplete the cell's mechanisms of dealing with oxidative compounds that can then activate the body's regulatory response (Cooper and Hanigan 2018). If this response is activated in a homeostatic or small amount, the body can successfully cope with these reactive species. However, if it is continuously or highly activated, it can lead to cell and tissue damage. Additionally, this is the general overview of the action of chemicals on their targets. Traditionally, it was thought that chemicals act on their intended primary target by an effector pathway, leading to therapeutic effects (Yuan et al. 2018). Any off-target effects, which occur when the chemical acts on another target, activate a different effector pathway. However, recent research shows that this is not always the case (Cruz-Migoni et al. 2019). Understanding the mechanisms of toxicology is key to predicting the risk of adverse chemical reactions in patients. For example, when it comes to liver toxicity, it is important to be able to gauge the concentration of the chemical at the hepatocyte site at the liver cell, as well as measure different types of dynamic biomarkers such as transaminase in blood samples. Additionally, the understanding of the dynamics of enzymes in the liver can help to predict chemical-induced liver injury (Yu et al. 2018). When it comes to the heart, the action of chemicals on different ion channels in the cardiomyocyte can be assessed, and the dynamic biomarkers, such as changes in the acuity of the electrocardiogram or different types of depolarizations, can be used to predict arrhythmias. Combining different modeling approaches enables better predictions for chemical toxicity related with cardiac or lung anomalies via liver metabolic profiling of chemicals. For example, human ether-à-go-go-related gene (hERG) channel profiler is included in the toxicological categorization of many hERG QSAR models (Seierstad and Agrafiotis 2006).

**Table 1** List of ADMET profiling tools (Shin et al. 2017)

| Endpoints                    | Type | Program   | Service |
|------------------------------|------|---|---------|
| <b>Absorption</b>            |      |   |         |
| PAMPA                        | SW   | MolCode toolbox   | C       |
| caco-2                       | SW   | preADMET, MolCode toolbox, Discovery Studio, volsurf+, QikProp  | W       |
|                              | DB   | the ADME database   | W       |
| MDCK                         | SW   | preADMET, QikProp, ADMETpredictor   | C       |
|                              |      | preADMET  | W       |
| HIA                          | SW   | ACD/percepta, preADMET, MolCode toolbox, Discovery Studio, ADMEWORKS Predictor, Stardrop                                    | C       |
|                              |      | PK/DB, admetSAR, preADMET   | W       |
|                              | DB   | the ADME database, PK/DB  | W       |
| Skin permeability            | SW   | preADMET, ADMETpredictor  | C       |
|                              |      | preADMET  | W       |
| Corneal permeability         | SW   | ADMETpredictor  | C       |
|                              | BA   | ACD/percepta, Impact-F  | C       |
|                              |      | PK/DB   | W       |
|                              | DB   | Pact-F  | C       |
|                              |      | the ADME database, PK/DB  | W       |
| <b>Distribution</b>          |      |   |         |
| BBB                          | SW   | ACD/percepta, preADMET, MolCode toolbox, Discovery Studio, ADMEWORKS Predictor, volsurf+, Stardrop, QikProp, ADMETpredictor | C       |
|                              |      | PK/DB, admetSAR, preADMET   | W       |
|                              | DB   | the ADME database, Chembench, PK/DB   | W       |
| VD                           | SW   | ACD/percepta, volsurf+, ADMETpredictor  | C       |
|                              | DB   | PK/DB   | W       |
| PPB                          | SW   | ACD/percepta, preADMET, MolCode toolbox, Discovery Studio, volsurf+, stardrop, QikProp, ADMETpredictor                      | C       |
|                              |      | PK/DB   | W       |
|                              | DB   | Chembench, PK/DB  | W       |
| Partition coefficient        | SW   | MolCode toolbox, ADMETpredictor   | C       |
| <b>Metabolism/excretion</b>  |      |   |         |
| P-gP inhibition              | SW   | ACD/percepta, ADMETpredictor  | C       |
|                              |      | preADMET, PK/DB, admetSAR   | W       |
|                              | DB   | the ADME database   | W       |
| P-gP substrates              | SW   | ADMEWORKS Predictor   | C       |
|                              |      | Chembench, admetSAR   | W       |
|                              | DB   | Chembench   | W       |
| Regioselectivity phase 1     | SW   | QikProp   | C       |
|                              |      | FAME  | W       |
| Regioselectivity phase 2     | SW   | MEXAlert  | C       |
|                              |      | FAME  | W       |
| Phase II substrate/inhibitor | SW   | preMetabo   | W       |
| Metabolic stability          | SW   | volsurf+  | C       |

Such profiler are created using boundaries based on repeated dose toxicity test data pulled from the Hazard Evaluation Support System (HESS) database.

IPBPB models are at center stage to explore some of the current approaches for predicting the concentration of chemicals in different compartments of the body, such as the central components of PBPK models (Kuepfer et al.

2016). PBPK models are used to describe how chemicals are absorbed, distributed, metabolized and eliminated. Through the use of current or previous knowledge from literature, these models capture the underlying physiological and mechanistic components, with the ultimate goal of predicting the concentration of chemicals in the plasma and the site of action (Abouir et al. 2021). A PBPK model looks

like a main compartment that consists of different blood circulation systems coupled to different compartments that represent the organs or tissues. Each compartment can be further specified and more information can be added in order to capture more mechanistic insight.

We can identify models associated to different types of toxicity, but the location of genes in these networks does not necessarily give us more insight into the mechanisms of regulation (Cordes et al. 2018). To gain a better understanding of the effector pathways, we need to incorporate additional data and use approaches such as linear programming algorithms or network-based approaches. We can also leverage on the wealth of omics data that has been generated for human and other organisms in order to construct the signaling layer. Mechanistic categorization schemes provide an organized way to identify key chemical properties based on published or expert knowledge (e.g. DNA binding by OASIS), helping to make informed decisions (Neuwoehner et al. 2008). Modeling chemical responses is of great importance, and the incorporating chemicals transformation into model has to be a key factor in metabolic modeling. In recent years, there has been a surge of research into genome-scale metabolic networks, which are used to simulate human metabolism at various levels, including cell, tissue, and organ. These networks are vast, with hundreds of coupled ordinary differential equations and thousands of metabolites and reactions, as seen in a model of the parasite and the comprehensive model of human metabolic (Carey et al. 2022). Developing these models requires immense effort, and often the collaboration of multiple institutions.

Therefore, how are metabolic models developed? This process starts by studying the enzymes and genes involved in metabolic reactions, using genomic, transcriptomic, and proteomic data to figure out if a particular protein is expressed in the tissue of interest (Wang and Zhang 2014). Then, all relevant literature is reviewed to identify the reactions taking place. After this information is represented in a set of equations, the model is simulated by constraint-based approaches to check if it matches the data. Since some of the reactions lack the necessary parameters, constraint-based approaches such as flux balance analysis are used to obtain quantitative analysis of metabolic flux at steady state (Dai and Locasale 2017).

### **Expanding genome-scale metabolic network (GSMN) with structural information**

How can we improve our metabolic model? Currently, there is a lot of research being done in this area. For example, adding information on protein structure and enzyme conformation can help us understand the impact of genetic variation. This includes knowledge on pharmacogenomics and

genomics, which may influence toxic response of a chemical (Gu et al. 2019). Additionally, looking at the three-dimensional structure of proteins and incorporating information on sequence variations associated with chemical response and disease can help build a knowledge base. Finally, molecular dynamic simulations can be used to predict the effect of mutations in the protein structure (Singh et al. 2022b) and its function, particularly its binding to certain chemicals or drugs (Hirano and Kameda 2021). In this section, we will briefly discuss through the process of combining PBPK modeling approaches, genomic scale metabolic network, and a model that regulates gene expression of one of the key enzymes related to metabolism of a chemical. PBPK model can be used to predict the concentration of toxic metabolites in the liver in the presence of different perturbations, such as chronic stress in people exposed to certain toxic environment (Maldonado et al. 2017). Through this approach, one can calculate the chemical metabolized in the whole body while considering the metabolism in the liver and gene expression of key enzymes.

An advantage of QSAR and PBPK approaches is the ability to incorporate variability in chemical action into the models, though it is not clear how this is done (Knaak et al. 2012). However, open question remains that should compound toxicologists adjust the models according to individual patient data, or is the variability already included in the models as they are in current state. Structural information of a protein can be used to model the variation of its response to chemical use. This can be done by different approaches, such as building a population of models by changing different parameters (3D state of amino acids, alpha helices, beta sheets, etc.). Additionally, it is possible to include food interactions and stress in metabolism models by considering their effects at the gene regulation level (Yau and Potenza 2013). For example, an increase in cortisol impacts the synthesis of a particular enzyme, which is explained by a gene regulatory network (Simmonds et al. 1984).

### **Nano-quantitative structure activity relationship (nano-QSAR)**

There are different methods in computational modeling such as quantitative structure activity relationship (QSAR)/ nano-QSAR, read across and data-driven profiling. The nano-QSAR or nano-QNTR (where N-nanostructure and T-toxicity) or nano-QNAR (where N-nanostructure and A-activity) approach can be useful in predicting the toxic potential of nanomaterials. The nano-QSAR approach statistically establish relationship between independent variables (physicochemical properties) and dependent variables (toxic effects) (Singh et al. 2023). In the past 2.1 decades,



the research on QSAR and nanoparticles has grown tremendously especially after year 2012 showing the importance of computational modeling in field of nanotoxicology. To interpret the research trends investigation, four ways can be used, namely

- Cumulative curve
- Density visualization map of co-occurrence of keywords
- Thematic map
- Conceptual structure map and keyword clusters

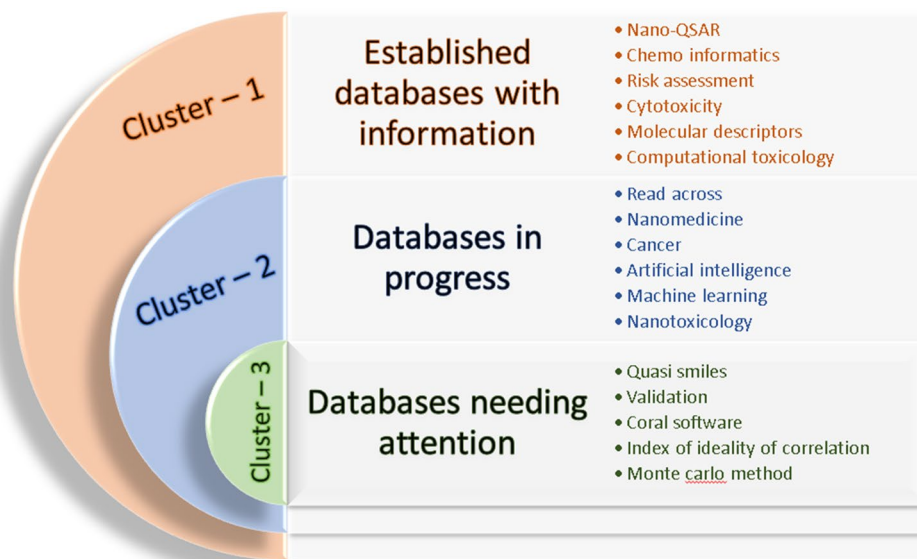
The Density Visualization map involves a yellow–green–blue color scheme which reflects the hotspots of nano-QSAR research (Singh et al. 2019). These color schemes are in order of the decreasing item densities like yellow color represent hot research area (e.g.: QSAR, validation, cytotoxicity, etc.), whereas blue color represent the opposite. It is done via “VOS viewer” software (VOSviewer 2022). The Thematic map contains four quadrants namely motor theme, basic theme, emerging/declining theme, specialized/niche theme. The motor theme involve topics that have well-developed and important themes. It involves keywords like drug-delivery, descriptor selection and design (Di Cosmo et al. 2021). In basic theme, the topics important for research field but not developed are included, such as keywords optimal descriptors, cytotoxicity, prediction and toxicity. In Conceptual Structure map (Bibliographic clustering analysis), the themes are grouped into three clusters represented with red, blue and green colors including drug discovery; engineered nanomaterials; and correlation and logic (CORAL software), respectively. The clusters with large number of topics are considered to be saturated

with research areas, while the less dense ones require more research or has scope for future (Fig. 5).

The nanoscale quantitative structure–activity relationship (nano-QSAR) is a computational technique which helps to understand the relationship between physical and chemical properties of nanomaterials and their biological effect on living organisms (Fortino et al. 2022). It predicts the biological activity of nanomaterials using quantum mechanics and statistical analysis. In nano-QSAR modeling, the mathematical relationship between variance in molecular properties (descriptors) and the variance in biological activity is obtained (Mikolajczyk et al. 2018). In nano-QSAR modeling, the primary objective is collection of data and data preprocessing.

The databases are obtained from various sources such as literature, databases, experiments and integrated sources. Under data collection, three main task comes into play—database, identify descriptors and endpoint selection (Singh et al. 2021a). The quality of data used determines the output of your assessment. Therefore, it is better to use various sources (good quality) rather than limited information on standardized protocol (poor quality) to get a reliable data output. The quality of nanoparticle data can be evaluated in several ways. One approach is to assess the accuracy and precision of the data, as well as the methods used to collect and analyze the data. Another approach is to evaluate the relevance and completeness of the data, and whether it is sufficient to answer the research questions at hand. Additionally, it is important to consider the credibility of the sources of the data and the expertise of the researchers involved in the study (Ballow et al. 1998). Ultimately, the quality of nanoparticle data can be determined by how well it meets the needs of the research and

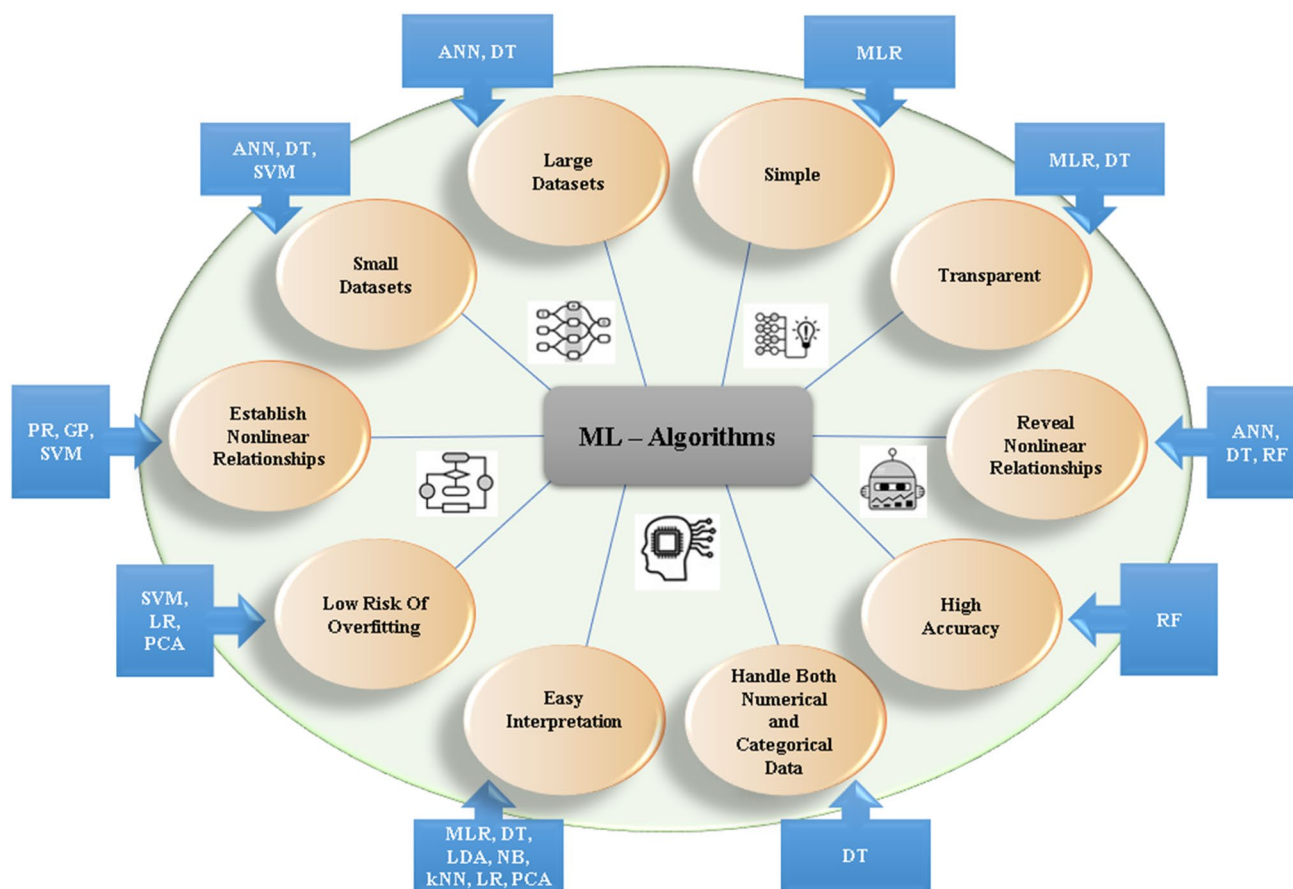
**Fig. 5** The conceptual structure map or bibliographic clustering analysis represented in three zones—red, blue and green cluster (color figure online)



how well it stands up to scrutiny from other experts in the field. To evaluate the quality of obtained nanoparticle data referred as ‘Nanosecurity’, the criteria of “FAIR” principles should be fulfilled where “FAIR” stands for findable, accessible, interoperable and reusable, respectively (Ammar et al. 2020).

The typical QSAR/QSPR approach is assumed to be generated on complete, homogeneous data, which are obtained in the same conditions. Unfortunately, when nanomaterial characteristics are mostly partial or performed in varied conditions, it is hard to include this information in typical modeling. To overcome this issue, an approach called perturbation approach is used that merges different kind of experimental data independent of their measurement conditions by identifying the problem and then adding small variation term to predict solution. Hence, a combined nano-QSAR perturbation approach can help predict the toxicity of nanoparticles under different experimental conditions with better results (Wyrzykowska et al. 2019).

In a conventional nano-QSAR model, we could predict results with only one endpoint and, therefore, have to create multiple QSAR equations for each end points. What if we could include multiple endpoints in a single model equation? It can become more practical, robust, reliable and economical. This can be achievable using a multitarget or multitasking QSAR approach (mt-QSAR), the Box-Jenkins moving average method-based software “QSAR-Co” (QSAR-Co 2021) or an advanced python-based toolkit QSAR-Co-X (Ambure et al. 2019; Halder and Dias Soeiro Cordeiro 2021; QSAR-Co-X 2023). It uses a single QSAR model equation to predict end points with different experimental or theoretical conditions and different biological targets. As shown in Fig. 6, the upgraded version has advantages like high reproducibility of linear modeling; automatically perform the diagnosis of inter-collinearity among variables; reduced computation time by keeping only random division for dataset division; automatically generate validation set and calculate



**Fig. 6** Advantages of different ML algorithms. For example, ANN is useful for both small and large datasets. SVM is suited for non-linear relationship and avoids overfitting. DT is versatile for transparency,

empirical and categorical data. RF better fits with non-linear relationship with high accuracy

its statistical parameters; more number of Box-Jenkins operators; availability of Yc randomization to incorporate the influence of experimental elements; several non-linear modeling tools (kNN, SVM, RF, NB, GB and MLP); comparative analysis of multiple machine learning methods; hyperparameter tuning options for machine learning methods; and condition-wise prediction to understand how the developed model performs against individual experimental conditions, particularly for large datasets (Halder and Dias Soeiro Cordeiro 2021).

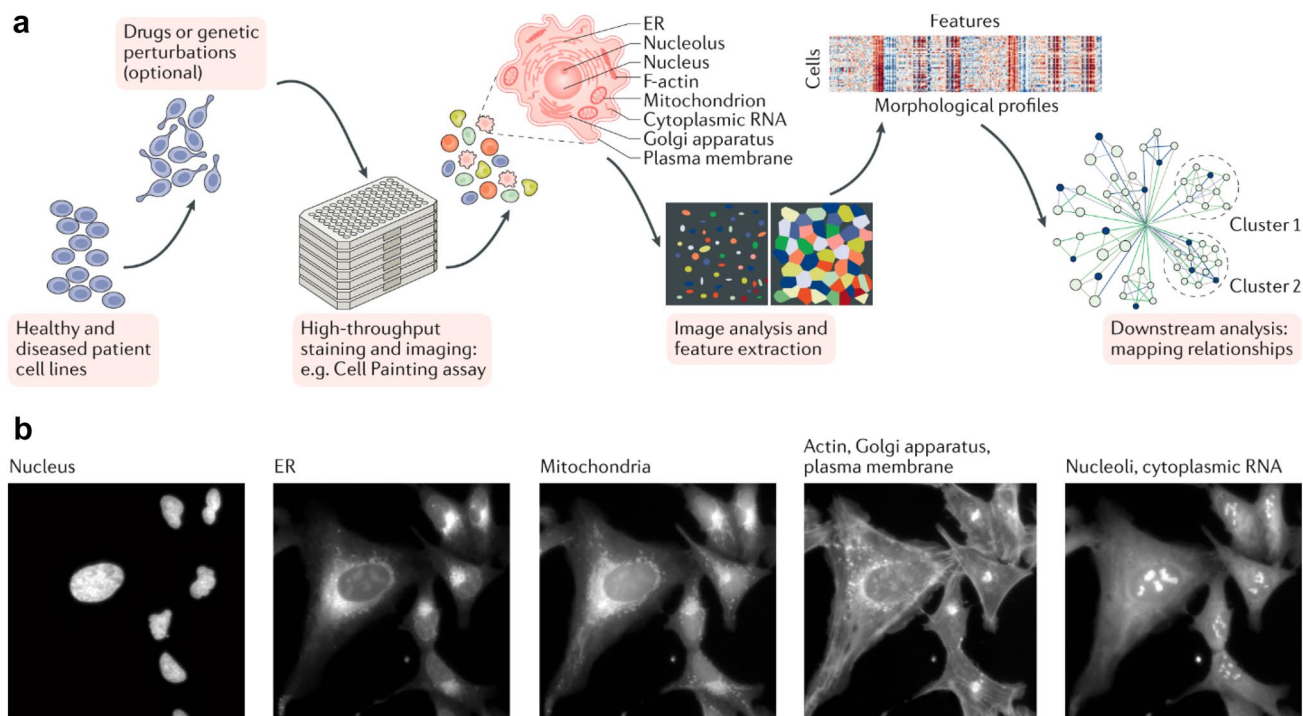
The applicability domain of a QSAR model is the response and chemical structure space in which the model makes predictions with a given reliability. According to OECD guidelines, applicability domain (AD) reflects the fact that QSARs are unavoidably associated with limitations in terms of the types of chemical structures, physico-chemical properties, and mechanisms of action for which the models can generate reliable predictions (Maharjan et al. 2022). If a new compound exists in the AD of the developed model, only then can the developed model predict the compound precisely. It is extremely useful for QSAR developers to have information about the AD of the developed model to identify interpolation (true predictions) or extrapolation (less reliable predictions) (Veerasamy et al. 2011). The applicability domain can be defined using different methods such as the value range of the variables, value ranges of principal components of variables, optimal prediction space (TPKAT), geometric methods, probabilistic density distribution methods, distance-based methods, etc. The most common method is distance-based method (DM), defined as any numerical measure of the prediction uncertainty for a given compound by the model or measures the reliability of predictions. According to DM, the model is under AD if the distance from the molecule to the distance of the training set is lower than a predefined threshold. A new software to calculate AD is “Model disturbance index tool” available on NanoBRIDGES project website (NanoBridges; Ruiz and Gómez-Nieto 2018).

Model interpretation is a way to comprehend and provide an explanation for the variables that cause the model to produce a response function. There are two main approaches in model interpretation that is, machine learning (ML)-dependent and ML-agnostic. The ML-dependent model interpretation uses regression coefficients, rule extraction, layer-wise relevance propagation (LRP), CAM and GRAD-CAM. The ML-agnostic approach is based on sensitivity analysis, partial derivatives, feature importance by perturbation, integrated gradients and Shapley sampling values (Matveieva and Polishchuk 2021).

## High content image-based screening data and toxicological databases

The high-content image-based screening (HCIBS) is a type of phenotypic drug discovery approach where biological images are used to analyze cells and tissues specific compounds or information (Singh et al. 2022a). HCIBS data typically include images of cells or tissues, as well as quantitative data on the various cellular parameters being measured (Fig. 7). The data may be collected from cells or tissues that have been treated with different compounds or conditions, in order to identify changes in cellular phenotype or gene expression. HCIBS data can be used to inform the development of predictive models to improve the accuracy of toxicological predictions (Antoniou et al. 2019). This method involves automated imaging techniques that captures high-resolution images of the cells and these can be analyzed using image analysis software to extract quantitative data. This includes imaging technique such as fluorescence microscopy, confocal microscopy, atomic force microscopy, etc. (Chandrasekaran et al. 2021; Lin et al. 2020).

In today’s world of internet, there are many sources available to gather an information where data are huge and scattered. An effective platform is required to gather relevant data at one place and deliver an integrated accurate information. The big data can be characterized by attributes like high volume of data (high-content screening/HCS), speed of data generation (high-throughput screening/HTS), different types of data (omics, chemical structures), variability (genetic/population variations), validity (specific endpoint), visibility (access to data sources) and adequacy for specific goal. Although the data on toxicology is huge but its concern over sharing, accessibility, processing, quality, comparability, interoperability, integration and relevancy limits the applicability in predictive toxicology (Richarz 2019). It is possible to combine data from various sources and exposure from various media and chemical sectors (for example, collected within the framework of various legislations) at various times, and then integrate the data to provide an overall big data resource that is closer to the diverse real-life exposure to chemicals/chemical mixtures and for evaluating co-exposures. For a better evaluation of mixtures, data can be mined to identify trends and clarify the mechanisms underlying chemical interactions. There are many toxicological databases available online shown in the following table with their use and websites (Table 2) (Ji et al. 2021; Pawar et al. 2019).



**Fig. 7** **a** An overview of typical steps in the workflow for generating image-based profiles from biological samples. **b** Example images from the Cell Painting assay often used for image-based profiling (Chandrasekaran et al. 2021)

## Challenges and future perspectives

Nano-QSAR modeling presents several challenges that can make it difficult to develop accurate and reliable models. Some of these challenges include:

- **The complex and multi-dimensional nature of nanomaterials:** The current list of descriptors is not enough to accurately predict the toxicity of nanoparticles due to high complexity and diversity of nanostructures. Thus, we need to find out some nano-specific descriptors that are most relevant to the activity of particular nanomaterial to get an accurate prediction of nano-toxicity.
- **The lack of high-quality experimental data:** In order to develop a reliable QSAR model, it is necessary to have a large and diverse dataset of experimental data. However, experimental data on the activity of nanomaterials are often limited, which can make it difficult to develop accurate models.
- **The lack of standardization in the field:** There are currently no widely accepted standards for describing the structures of nanomaterials, which can make it difficult to compare and evaluate different models.

- **The potential for overfitting:** Overfitting is a common problem in machine learning, where a model becomes too closely matched to the specific data used to train it, and thus performs poorly on new data. Overfitting can be a particular concern in nano-QSAR modeling, due to the limited availability of experimental data.
- **The ability of nanoparticles to dynamically interact with the exposure conditions and trigger series of biological effect makes it even more difficult to predict the toxic potential of nanoparticle in different environment such as change in hydrophobic interactions, hydrogen bonding, corona formation with plasma and serum components.**

However, there are some challenges specific to nano-QSAR modeling, including the complexity and variability of the structural properties of nanomaterials, as well as the lack of standardized methods for measuring their biological activity. As a result, the development of accurate and reliable nano-QSAR models can be a challenging task. To overcome these challenges, it is important to carefully select the training dataset and the descriptors used to represent the structural properties of the nanomaterials. The use of some new computational strategies

**Table 2** Different databases used in toxicology

| Database  | Usage  | Website   |
|---|--|---|
| Toxline (Toxnet)  | Provide literature on biochemical, pharmacological, physiological, and toxicological effects of drugs and other chemicals  | <a href="https://toxnet.nlm.nih.gov/">https://toxnet.nlm.nih.gov/</a>   |
| Aggregated Computational Toxicology Online Resource (AcToR)       | Include chemical structure, physicochemical values, in vitro assay data and in vivo toxicology data  | –   |
| Gene-Tox database   | Comprises mutagenicity data for more than 3,000 chemicals  | <a href="https://www.nlm.nih.gov/databases/download/genetox.html">https://www.nlm.nih.gov/databases/download/genetox.html</a>   |
| Toxicity Reference Database (ToxRefDB)                            | contains in vivo study data from over 5900 guideline or guideline-like studies for over 1100 chemicals   | –   |
| ChemBL  | Brings together chemical, bioactivity and genomic data   | <a href="https://www.ebi.ac.uk/chembl/target_report_card/CHEMBL1245/">https://www.ebi.ac.uk/chembl/target_report_card/CHEMBL1245/</a>   |
| Toxicity Value Database (Tox ValDB)                               | Includes data on thousands of chemicals with an emphasis on quantitative estimates of relevant points-of-departure from in vivo toxicology studies, such as no- and low-observable adverse effect levels, screening levels, reference doses, tolerable daily intake, etc | <a href="https://www.epa.gov/chemical-research/downloadable-computational-toxicology-data#:~:text=The%20Toxicity%20Value%20Database%20(Tox%20ValDB,screening%20levels%2C%20reference%20doses%2C%20tolerable">https://www.epa.gov/chemical-research/downloadable-computational-toxicology-data#:~:text=The%20Toxicity%20Value%20Database%20(Tox%20ValDB,screening%20levels%2C%20reference%20doses%2C%20tolerable</a> |
| RepDose and FedTex database                                       | Useful sources for No Observed (Adverse) Effect Level (NO(A)EL) or Lowest Observed (Adverse) Effect Level (LO(A)EL) values from repeated dose studies for reproductive and developmental toxicity endpoints  | <a href="http://cefic-iri.org/toolbox/fedtex/">http://cefic-iri.org/toolbox/fedtex/</a>   |
| HESS database   | Information on 28 day repeat dose toxicity studies for 289 industrial chemicals in rats and includes additional rat metabolism datasets and information on ADME in rats and humans   | <a href="https://www.nite.go.jp/en/chem/qsar/hess-e.html">https://www.nite.go.jp/en/chem/qsar/hess-e.html</a>   |
| ECOTOX  | Provides information on adverse effects of single chemical stressors to ecologically relevant aquatic and terrestrial species  | <a href="https://cfpub.epa.gov/ecotox/">https://cfpub.epa.gov/ecotox/</a>   |
| T3DB (Toxin and Toxin Target database or Toxic Exposome database) | It combines detailed toxin data with comprehensive toxin target information  | <a href="http://www.t3db.ca/">http://www.t3db.ca/</a>   |
| Comparative Toxicogenomics Database (CTD)                         | Includes more than 30.5 million toxicogenomic connections relating chemicals/drugs, genes/proteins, diseases, taxa, Gene Ontology (GO) annotations, pathways, and gene interaction modules   | <a href="https://ctdbase.org/">https://ctdbase.org/</a>   |
| Open-TG-GATES   | Stores gene expression profiles and traditional toxicological data derived from in vivo (rat) and in vitro (primary rat hepatocytes, primary human hepatocytes) exposure to 170 compounds at multiple dosages and time points  | <a href="https://dbarchive.biosciencecd.bc.jp/en/open-tggates/download.html">https://dbarchive.biosciencecd.bc.jp/en/open-tggates/download.html</a>   |
| Toxygates   | New interactive version of data from Open-TG-GATES covering 24,011 samples and 170 compounds   | <a href="https://toxygates.nibiohn.go.jp/">https://toxygates.nibiohn.go.jp/</a>   |
| ChemTunes ToxGPS  | Consists of in vitro and in vivo toxicity endpoint specific alerting chemotypes; mechanism of action (MOA) based QSAR models, weight of evidence (WoE) outcomes, and ToxGPS datasets   | <a href="https://chemtunes.com/">https://chemtunes.com/</a>   |

Table 2 (continued)

| Database   | Usage   | Website   |
|--|---|---|
| Nanoparticle Information Library                   | Share information about nanomaterials, including their health and safety-related properties   | <a href="http://nanoparticlelibrary.net/">http://nanoparticlelibrary.net/</a>   |
| Cancer Nanotechnology Laboratory/caNano Lab        | Expedite and validate the use of nanotechnology in biomedicine  | <a href="https://cananolab.nci.nih.gov/">https://cananolab.nci.nih.gov/</a>   |
| DaNa Knowledgebase                                 | Help understanding the impacts of nanomaterials for humans and the environment  | <a href="https://www.nanopartikel.info/en/">https://www.nanopartikel.info/en/</a>   |
| Nanomaterial registry                              | Help understanding the fundamental properties of nanomaterials  | <a href="http://nanohub.org/">http://nanohub.org/</a>   |
| eNanomapper  | Develop a computational framework for nano-toxicity data management   | <a href="http://www.enanomapper.net/">http://www.enanomapper.net/</a>   |
| S <sup>2</sup> NANO                                | Develop and commercialize safe and sustainable nano-products  | <a href="http://portal.s2nano.org/">http://portal.s2nano.org/</a>   |
| NanoSolveIT  | A new concept, fingerprinting of nanomaterials, is proposed. Predicting the properties, functions, and hazards of nanomaterials by developing and integrating nanoinformatics | <a href="https://nanosolveit.eu/">https://nanosolveit.eu/</a>   |
| EU NanoSafety Cluster                              | Study aspects include toxicology, ecotoxicology, exposure assessment, mechanisms of interaction, risk assessment, and standardization   | <a href="https://www.nanosafetycluster/">https://www.nanosafetycluster/</a>   |
| NanoWerks Nanomaterial Database                    | Help the nanotechnology community to research nanomaterials   | <a href="https://www.nanowerk.com/">https://www.nanowerk.com/</a>   |
| Nanomaterial-Biological Interactions Knowledgebase | Help understanding the mechanism of nanomaterial exposure effects in biological systems   | <a href="http://nbi.oregonstate.edu/">http://nbi.oregonstate.edu/</a>   |
| InterNano  | Provides information about nanomanufacturing research, government reports on nanomanufacturing  | <a href="http://www.internano.org/">http://www.internano.org/</a>   |
| NanoHUB  | A web-based infrastructure for e-collaboration in the nanotechnology simulation community   | <a href="https://nanohub.org/resources/">https://nanohub.org/resources/</a>   |
| Nanotechnology Characterization Laboratory         | Performs and standardizes the preclinical characterization of nanomaterials intended as cancer therapeutics   | <a href="https://ncl.cancer.gov/">https://ncl.cancer.gov/</a>   |
| NanoMILE   | Contain characterization data and high-throughput screening toxicity data of nanomaterials  | <a href="https://ssl.biomax.de/nanomile/cgi/login_bioxm_portal.cgi">https://ssl.biomax.de/nanomile/cgi/login_bioxm_portal.cgi</a> |
| NanoDatabank                                       | Design with simplicity of nanomaterial data storing and sharing   | <a href="http://nanoinfo.org/nanodatabank/">http://nanoinfo.org/nanodatabank/</a>   |
| PubVINAS   | —   | <a href="http://www.pubvinas.com/">http://www.pubvinas.com/</a>   |

to make diverse data more inclusive can help deal with shortage of homogenous experimental data. Additionally, the use of advanced machine learning algorithms and appropriate validation techniques can also improve the accuracy and reliability of nano-QSAR models.

**Acknowledgements** AVS Thanks BfR for internal Grant (award No. SFP-1322-735).

**Author contributions** Conceptualization, MV and AVS; writing—original draft preparation, MV and AVS; writing—review and editing, VKD, PK, PL and SC, AKD and BN; Graphic Design, AVS, NG, AG and MV; funding acquisition: PL, AVS and AL. All authors have read and agreed to the published version of the manuscript.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Data availability** The data that support the findings of this review are available in the references cited in the text. Additional data related to this review are available from the corresponding author upon reasonable request.

## Declarations

**Conflict of interests** Authors state no conflict of interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Abouir K, Samer CF, Gloor Y, Desmeules JA, Daali Y (2021) Reviewing data integrated for PBPK model development to predict metabolic drug–drug interactions: shifting perspectives and emerging trends. *Front Pharmacol*. <https://doi.org/10.3389/fphar.2021.708299>
- Ahamed N (2014) Ecotoxicity concert of nano zero-valent iron particles—a review. *J Crit Rev* 1(1):36–39
- Ambure P, Halder AK, González Díaz H, Cordeiro MNDS (2019) QSAR-Co: an open source software for developing robust multitasking or multitarget classification-based QSAR models. *J Chem Inf Model* 59(6):2538–2544. <https://doi.org/10.1021/acs.jcim.9b00295>
- Ammar A, Bonaretti S, Winckers L et al (2020) A semi-automated workflow for FAIR maturity indicators in the life sciences. *Nanomaterials* 10(10):2068
- Antoniou AN, Powis SJ, Kriston-Vizi J (2019) High-content screening image dataset and quantitative image analysis of Salmonella infected human cells. *BMC Res Notes* 12(1):1–4
- Asati S, Sahu A, Jain A (2021) Nanotoxicity: the dark side of nanoformulations. *Curr Nanotoxicity Prevent (discontinued)* 1(1):6–25. <https://doi.org/10.2174/2665980801999201230095324>
- Ballou CH, Amsden GW, Highet VS, Forrest A (1998) Pharmacokinetics of oral azithromycin in serum, urine, polymorphonuclear leucocytes and inflammatory vs non-inflammatory skin blisters in healthy volunteers. *Clin Drug Invest* 15(2):159–167
- Bayda S, Adeel M, Tuccinardi T, Cordani M, Rizzolio F (2019) The history of nanoscience and nanotechnology: from chemical–physical applications to nanomedicine. *Molecules* 25(1):112
- Carey MA, Medlock GL, Stolarczyk M, Petri WA Jr, Guler JL, Papin JA (2022) Comparative analyses of parasites with a comprehensive database of genome-scale metabolic models. *PLoS Comput Biol* 18(2):e1009870
- Chandrasekar V, Singh AV, Maharjan RS et al (2022) Perspectives on the technological aspects and biomedical applications of virus-like particles/nanoparticles in reproductive biology: insights on the medicinal and toxicological outlook. *Adv NanoBiomed Res* 2(8):2200010. <https://doi.org/10.1002/anbr.202200010>
- Chandrasekaran SN, Ceulemans H, Boyd JD, Carpenter AE (2021) Image-based profiling for drug discovery: due for a machine-learning upgrade? *Nat Rev Drug Discov* 20(2):145–159. <https://doi.org/10.1038/s41573-020-00117-w>
- Cooper AJL, Hanigan MH (2018) 10.17 - Metabolism of glutathione S-conjugates: multiple pathways. In: McQueen CA (ed) *Comprehensive toxicology*, 3rd edn. Elsevier, Oxford, pp 363–406
- Cordes H, Thiel C, Baier V, Blank LM, Kuepfer L (2018) Integration of genome-scale metabolic networks into whole-body PBPK models shows phenotype-specific cases of drug-induced metabolic perturbation. *NPJ Syst Biol Appl* 4(1):10. <https://doi.org/10.1038/s41540-018-0048-1>
- Cruz-Migoni A, Canning P, Quevedo CE et al (2019) Structure-based development of new RAS-effector inhibitors from a combination of active and inactive RAS-binding compounds. *Proc Natl Acad Sci* 116(7):2545–2550. <https://doi.org/10.1073/pnas.1811360116>
- Dai Z, Locasale JW (2017) Understanding metabolism with flux analysis: From theory to application. *Metab Eng* 43:94–102. <https://doi.org/10.1016/j.ymben.2016.09.005>
- Daoud NE-H, Borah P, Deb PK et al (2021) ADMET profiling in drug discovery and development: perspectives of in silico, in vitro and integrated approaches. *Curr Drug Metab* 22(7):503–522
- Di Cosmo A, Pinelli C, Scandurra A, Aria M, Daniello B (2021) Research trends in octopus biological studies. *Animals* 11(6):1808
- Domingues C, Santos A, Alvarez-Lorenzo C et al (2022) Where is nano today and where is it headed? A review of nanomedicine and the dilemma of nanotoxicology. *ACS Nano* 16(7):9994–10041. <https://doi.org/10.1021/acsnano.2c00128>
- Fortino V, Kinaret PAS, Fratello M et al (2022) Biomarkers of nanomaterials hazard from multi-layer data. *Nat Commun* 13(1):1–10
- Fuelle H-J, Lanctin V (2022) Regulatory affairs—communicating with health authorities principles of biomedical sciences and industry. Pp 177–202
- Granath B, Jalkestén E, Cotgreave I (2014) Strategic focus on 3R principles reveals major reductions in the use of animals in pharmaceutical toxicity. *PLoS ONE* 9(7):e101638
- Grandjean P (2016) Paracelsus revisited: the dose concept in a complex world. *Basic Clin Pharmacol Toxicol* 119(2):126–132. <https://doi.org/10.1111/bcpt.12622>
- Gu C, Kim GB, Kim WJ, Kim HU, Lee SY (2019) Current status and applications of genome-scale metabolic models. *Genome Biol* 20(1):1–18
- Gupta R, Srivastava D, Sahu M, Tiwari S, Ambasta RK, Kumar P (2021) Artificial intelligence to deep learning: machine intelligence approach for drug discovery. *Mol Diversity* 25(3):1315–1360. <https://doi.org/10.1007/s11030-021-10217-3>

- Halder AK, Dias Soeiro Cordeiro MN (2021) QSAR-Co-X: an open source toolkit for multitarget QSAR modelling. *J Cheminform* 13(1):29. <https://doi.org/10.1186/s13321-021-00508-0>
- Hemmerich J, Ecker GF (2020) In silico toxicology: from structure–activity relationships towards deep learning and adverse outcome pathways. *Wires Comput Mol Sci* 10(4):e1475. <https://doi.org/10.1002/wcms.1475>
- Hirano A, Kameda T (2021) Aromaphilicity index of amino acids: molecular dynamics simulations of the protein binding affinity for carbon nanomaterials. *ACS Appl Nano Mater* 4(3):2486–2495. <https://doi.org/10.1021/acsnm.0c03047>
- Ji Z, Guo W, Sakkiah S, Liu J, Patterson TA, Hong H (2021) Nanomaterial databases: Data sources for promoting design and risk assessment of nanomaterials. *Nanomaterials* 11(6):1599
- Kim JH, Scialli AR (2011) Thalidomide: the tragedy of birth defects and the effective treatment of disease. *Toxicol Sci* 122(1):1–6. <https://doi.org/10.1093/toxsci/kfr088>
- Knaak JB, Dary CC, Zhang X et al (2012) Parameters for pyrethroid insecticide QSAR and PBPK/PD models for human risk assessment. In: Whitacre DM (ed) *Reviews of environmental contamination and toxicology*. Springer, New York, pp 1–114
- Kuepfer L, Niederal C, Wendl T et al (2016) Applied concepts in PBPK modeling: how to build a PBPK/PD model. *CPT Pharmacometr Syst Pharmacol* 5(10):516–531. <https://doi.org/10.1002/psp4.12134>
- Kumar R, Saha P (2022) A review on artificial intelligence and machine learning to improve cancer management and drug discovery. *Int J Res Appl Sci Biotechnol* 9(3):149–156
- Lin S, Schorpp K, Rothenaigner I, Hadian K (2020) Image-based high-content screening in drug discovery. *Drug Discov Today* 25(8):1348–1361. <https://doi.org/10.1016/j.drudis.2020.06.001>
- Maharjan RS, Singh AV, Hanif J et al (2022) Investigation of the associations between a nanomaterial’s microrheology and toxicology. *ACS Omega* 7(16):13985–13997
- Maldonado EM, Leoncik V, Fisher CP, Moore JB, Plant NJ, Kierzek AM (2017) Integration of genome scale metabolic networks and gene regulation of metabolic enzymes with physiologically based pharmacokinetics. *CPT Pharmacometr Syst Pharmacol* 6(11):732–746. <https://doi.org/10.1002/psp4.12230>
- Matveieva M, Polishchuk P (2021) Benchmarks for interpretation of QSAR models. *J Cheminform* 13(1):41. <https://doi.org/10.1186/s13321-021-00519-x>
- Mikolajczyk A, Gajewicz A, Mulkiewicz E et al (2018) Nano-QSAR modeling for ecosafe design of heterogeneous TiO<sub>2</sub>-based photocatalysts. *Environ Sci Nano* 5(5):1150–1160. <https://doi.org/10.1039/C8EN00085A>
- NanoBridges (2023) <http://nanobridges.eu/>. Accessed Dec 2022
- Neuwoehner J, Junghans M, Koller M, Escher BI (2008) QSAR analysis and specific endpoints for classifying the physiological modes of action of biocides in synchronous green algae. *Aquatic Toxicol* 90(1):8–18. <https://doi.org/10.1016/j.aquatox.2008.07.010>
- Pawar G, Madden JC, Ebbrell D, Firman JW, Cronin MTD (2019) In silico toxicology data resources to support read-across and (Q) SAR. *Front Pharmacol*. <https://doi.org/10.3389/fphar.2019.00561>
- Pérez Santín E, Rodríguez Solana R, González García M et al (2021) Toxicity prediction based on artificial intelligence: a multidisciplinary overview. *Wiley Interdiscip Rev Comput Mol Sci* 11(5):e1516
- Petejova N, Martinek A, Zadrzil J, Teplan V (2019) Acute toxic kidney injury. *Ren Fail* 41(1):576–594. <https://doi.org/10.1080/0886022X.2019.1628780>
- QSAR-Co (2021) <https://sites.google.com/view/qsar-co>. Accessed Dec 2022
- QSAR-Co-X (2023) <https://github.com/ncordeirfcup/QSAR-Co-X>. Accessed Jan 2023
- Richarz A-N (2019) Big data in predictive toxicology: challenges, opportunities and perspectives.
- Ruiz IL, Gómez-Nieto M (2018) Study of the applicability domain of the QSAR classification models by means of the rivalry and modelability indexes. *Molecules*. <https://doi.org/10.3390/molecules23112756>
- Seierstad M, Agrafiotis DK (2006) A QSAR model of hERG binding using a large, diverse, and internally consistent training set. *Chem Biol Drug Design* 67(4):284–296. <https://doi.org/10.1111/j.1747-0285.2006.00379.x>
- Shin HK, Kang Y-M, No KT (2017) Predicting ADME properties of chemicals. In: Leszczynski J, Kaczmarek-Kedziera A, Puzyn T, Papadopoulos M, Reis H, Shukla M (eds) *Handbook of computational chemistry*. Springer International Publishing, Cham, pp 2265–2301
- Simmonds P, Miles J, Gerich J, Haymond M (1984) Increased proteolysis. An effect of increases in plasma cortisol within the physiological range. *J Clin Invest* 73:412–420
- Singh AV, Jahnke T, Xiao Y et al (2019) Peptide-induced biomineralization of tin oxide (SnO<sub>2</sub>) nanoparticles for antibacterial applications. *J Nanosci Nanotechnol* 19(9):5674–5686
- Singh AV, Ansari MHD, Rosenkranz D et al (2020a) Artificial intelligence and machine learning in computational nanotoxicology: unlocking and empowering nanomedicine. *Adv Healthc Mater* 9(17):1901862. <https://doi.org/10.1002/adhm.201901862>
- Singh AV, Maharjan R-S, Kanase A et al (2020b) Machine-learning-based approach to decode the influence of nanomaterial properties on their interaction with cells. *ACS Appl Mater Interfaces* 13(1):1943–1955
- Singh AV, Chandrasekar V, Janapareddy P et al (2021a) Emerging application of nanorobotics and artificial intelligence to cross the BBB: advances in design, controlled maneuvering, and targeting of the barriers. *ACS Chem Neurosci* 12(11):1835–1853
- Singh AV, Maharjan RS, Kromer C et al (2021b) Advances in smoking related in vitro inhalation toxicology: a perspective case of challenges and opportunities from progresses in lung-on-chip technologies. *Chem Res Toxicol* 34(9):1984–2002. <https://doi.org/10.1021/acs.chemrestox.1c00219>
- Singh AV, Romeo A, Scott K et al (2021c) Emerging technologies for in vitro inhalation toxicology. *Adv Healthc Mater* 10(18):2100633. <https://doi.org/10.1002/adhm.202100633>
- Singh AV, Chandrasekar V, Laux P et al (2022a) Micropatterned neurovascular interface to mimic the blood–brain barrier’s neurophysiology and micromechanical function: a BBB-on-CHIP model. *Cells* 11(18):2801
- Singh AV, Kayal A, Malik A et al (2022b) Interfacial water in the SARS spike protein: investigating the interaction with human ACE2 receptor and in vitro uptake in A549 cells. *Langmuir* 38(26):7976–7988
- Singh A, Maharjan R, Jungnickel H et al (2021) Evaluating particle emissions and toxicity of 3d pen printed filaments with metal nanoparticles as additives. *ACS Sustainable Chem Eng* 9:11724–11737
- Swaters D, van Veen A, van Meurs W, Turner JE, Ritskes-Hoitinga M (2022) A history of regulatory animal testing: what can we learn? *Alternatives Lab Anim* 50(5):322–329
- Veerasingh R, Rajak H, Jain A, Sivadasan S, Varghese CP, Agrawal RK (2011) Validation of QSAR models-strategies and importance. *Int J Drug Des Discov* 3:511–519
- VOSviewer (2022) proceed with citation. In: VOSviewer. <https://www.vosviewer.com/downloads/VOSviewer1.6.18.exe.zip>. Accessed Dec 2022
- Wang X, Zhang B (2014) Integrating genomic, transcriptomic, and interactome data to improve peptide and protein identification in shotgun proteomics. *J Proteome Res* 13(6):2715–2723. <https://doi.org/10.1021/pr500194t>



- Wyrzykowska E, Jagiello K, Rasulev B, Puzyn T (2019) Descriptors in nano-QSAR/QSPR modeling computational nanotoxicology. Jenny Stanford Publishing, pp 245–302
- Yau YH, Potenza MN (2013) Stress and eating behaviors. *Minerva Endocrinol* 38(3):255–267
- Yordanova D, Kuseva C, Tankova K et al (2019) Using metabolic information for categorization and read-across in the OECD QSAR Toolbox. *Comput Toxicol* 12:100102
- Yu K-N, Nadanaciva S, Rana P et al (2018) Prediction of metabolism-induced hepatotoxicity on three-dimensional hepatic cell culture and enzyme microarrays. *Arch Toxicol* 92(3):1295–1310. <https://doi.org/10.1007/s00204-017-2126-3>
- Yuan TL, Amzallag A, Bagni R et al (2018) Differential effector engagement by oncogenic KRAS. *Cell Rep* 22(7):1889–1902

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.