



**NovaMechanics**

Cheminformatics & Nanoinformatics Excellence

A study on valid *in silico* modelling tools and read-across approaches, including creation of case studies on read-across for specific (types of) nanomaterials

May 2023

## Table of Contents

Disclaimer .....	5
Abstract .....	5
Executive summary .....	6
Objective .....	6
Methodology .....	6
Results .....	7
Key findings .....	7
Literature review .....	7
Experts' opinion.....	10
Case studies.....	12
General conclusions .....	13
Outlook.....	13
1 Introduction .....	16
2 Main types of assessed <i>in silico</i> methodologies .....	18
2.1 nanoQSAR.....	18
2.2 Grouping and read-across .....	18
2.3 Physiologically-based pharmacokinetic modelling .....	19
2.4 Molecular dynamics simulations .....	20
2.5 Adverse outcome pathways .....	20
3 Extended literature search .....	21
3.1.1 Search strategy protocol .....	22
3.1.2 Selection of databases to be examined .....	23
3.1.3 Definition of search terms and Boolean operators .....	23
3.1.4 Resources' selection criteria.....	24
3.1.5 Synthesis, presentation, and interpretation of data.....	24
4 Information extraction .....	25
5 Critical analysis of the collected <i>in silico</i> tools and available models .....	27
5.1 nanoQSAR/read-across .....	31
5.1.1 Data and availability .....	31
5.1.2 Model development.....	36
5.1.3 Model validation.....	39
5.1.4 Applicability domain definition.....	46
5.1.5 Model dissemination .....	48
5.1.6 Quality assessment of the nanoQSAR and read-across studies.....	52
5.2 Physiologically-based pharmacokinetics models.....	56

5.3	Molecular dynamics simulations .....	58
5.4	Adverse outcome pathways .....	60
5.5	Other models.....	62
5.5.1	Causal Models .....	62
5.5.2	Exposure Models .....	63
5.5.3	Dose response .....	63
5.6	Models and methods implemented as user-friendly tools.....	63
6	Interviews with experts from Academia and Industry.....	65
6.1	Identification of the experts .....	65
6.2	Organisation, preparation, distribution of the questionnaires and collection of the data ....	66
6.3	Risk mitigation measures.....	67
6.4	Experts' information.....	68
6.5	<i>In silico</i> , grouping, and read-across methodologies .....	69
6.6	Data availability .....	72
6.7	Dissemination.....	74
6.8	Gaps and future steps .....	75
7	Case studies.....	77
7.1	Case study 1: Development of read-across models for the assessment of the anti-microbial activity of various carbon-based nanoparticles .....	77
7.1.1	Scope of the first case study.....	77
7.1.2	Results of the first case study .....	79
7.1.3	Discussion.....	86
7.2	Case study 2: Assessment of different tools for the prediction of TiO <sub>2</sub> -based NMs' endpoints 88	
7.2.1	Scope of the second case study.....	88
7.2.2	Results of the second case study.....	89
7.3	Case study 3: Combination of nanoinformatics models into SAPNets .....	115
7.3.1	Scope of the third case study .....	115
7.3.2	Results of the third case study .....	115
7.3.3	Discussion.....	120
8	Summary .....	124
8.1	Conclusions and outlook .....	125
	Abbreviations .....	128
	References.....	130
	Appendix I – List of assessed studies and models for nanoQSAR, grouping and read-across alternative methods .....	146

Appendix II – Nanoinformatics models and methodologies available via GUIs .....	176
A safe-by-design tool for functionalised nanomaterials: a web-service available through the Enalos Nanoinformatics Cloud Platform .....	176
Apellis .....	177
BioDaph .....	178
Cytotoxicity (cell viability) prediction for metal oxide NPs .....	180
DeepDaph .....	181
Ecotox models .....	181
Enalos QNAR iron oxide toxicity platform .....	182
Facet cytotoxicity prediction .....	183
GUIDEnano .....	184
INSIdE nano .....	185
Logistic regression model for the toxicity classification of SPIONs .....	186
Lung exposure dose calculator .....	188
Metal oxide nanoparticles cytotoxicity classification .....	190
MS <sup>3</sup> bD zeta potential predictive model .....	191
Nano-lazar .....	192
Nano Protein Corona model .....	193
NanoBio .....	194
Nanoinformatics model for zeta potential prediction powered by Enalos Cloud Platform .....	196
NanoInhale .....	197
NanoMixHamster .....	199
NanoPot .....	200
NanoProfiler .....	201
NanoSafer .....	202
NanoSerpA .....	202
NanoToxRadar .....	204
PBPK models and integration with the occupational exposure model .....	205
Prediction of MNPs uptake in PaCa2 cancer cells through Enalos Cloud Platform .....	206
QSAR for nano-mixtures .....	207
QsarDB models .....	208
Quantitative read-across (Read-Across-v4.1) .....	209
Random forest model for the genotoxicity classification of MWCNTs .....	212
RiskGONE <i>in vitro</i> dosimetry tool .....	213
Sbpot model .....	215
Tool for assessment of human exposure to nanomaterials (multi-box aerosol model) .....	215



toxFlow .....	218
Vythos .....	219
$\zeta$ -regression .....	221
Appendix III – Questionnaire on valid <i>in silico</i> modelling tools and read-across approaches.....	222
Preamble/Cover letter.....	222
Section 1 – Personal information .....	223
Section 2 – <i>In silico</i> , grouping, and read-across methodologies.....	223
Section 3 – Data availability.....	223
Section 4 – Dissemination .....	224
Section 5 – Gaps and future steps.....	224
Appendix IV – Graphs with the responses to the questionnaire .....	226
Section 1 – Personal information .....	226
Section 2 – <i>In silico</i> , grouping, and read-across methodologies.....	227
Section 3 – Data availability.....	229
Section 4 – Dissemination .....	232
Section 5 – Gaps and future steps.....	235
Appendix V – QMRF report for the EnaloskNN model (Case study 1) .....	237
Appendix VI – Quantitative measures of goodness-of-fit and predictivity for regression models .....	242
Correlation of determination ( $R^2$ ).....	242
External explained variance ( $Q^2_{\text{ext}}$ or $Q^2_{F1}$ ).....	242
Mean absolute error (MAE).....	242
Mean square error (MSE) .....	242
Root mean square error (RMSE) .....	242
Squared Pearson correlation coefficient ( $r^2$ ) .....	243
Golbraikh and Tropsha’s test.....	243

## Disclaimer

This publication is solely intended for information purposes and does not necessarily represent the official opinion of the European Chemicals Agency. The European Chemicals Agency is not responsible for the use that may be made of the information contained in this document.

## Abstract

The present study provides the European Union Observatory of Nanomaterials (EUON) and its stakeholders with insight on available frameworks and state-of-the-art (scientific) developments for read-across and other *in silico* approaches, which can be alternatives to conventional nanomaterials (NMs) hazard and risk assessments. In the course of this exercise, a systematic literature review was planned and carried out to identify the available models and tools for the computational assessment of NMs.

Based on the identified studies, a thorough analysis was performed to map the aspects of models that are properly implemented (model development, validation, and documentation), as well as any potential data gaps and flaws that could be improved in upcoming research. Next, expert opinions were collected and incorporated in the findings to present a complete and integrated view of the nanoinformatics field. Finally, the applicability, the strengths and limitations of the models were demonstrated through the development of three case studies using as many of the identified tools as possible.

Based on the collected results we concluded that the nanoinformatics field is moving forward, as there are already available reliable models and the infrastructure needed to implement them as user-friendly tools (e.g., supported by different EU-funded projects). Nonetheless, the main barriers and delays to the adoption of the *in silico* approaches by regulators continue to be the lack of transparency in the currently available methodologies (including algorithms, assumptions, validation process, and applicability domain definition), and of real-case examples/applications. Building trust in the use of appropriately reported and validated computational methods that are also scientifically supported, is therefore the next step toward *in silico* approaches gaining widespread public and regulatory acceptance as alternatives to conventional hazard and risk assessments.

## Executive summary

### Objective

Upon the request of the European Chemicals Agency (ECHA) and the European Union Observatory of Nanomaterials (EUON) NovaMechanics Ltd. has undertaken a project on the identification of valid *in silico* modelling tools (computational models) and read-across approaches, including the creation of case studies on read-across for specific (types of) nanomaterials (NMs).

The purpose of this report is to provide the ECHA/EUON and the interested groups with reliable and transparent information on the existing computational approaches for the risk assessment and safety-by-design (SbD) of novel NMs<sup>a</sup>. These approaches include nanoQSAR, grouping and read-across methods, adverse outcome pathways (AOPs), physiologically-based pharmacokinetic (PBPK) models and simulations (see §2). To address this mission, an extensive and systematic literature review was conducted to mine the existing tools and methodologies, as well as an online survey to collect the experts' opinion in the field of nanoinformatics. Furthermore, a representative number of the identified tools (10 out of 38) was tested to assess their functionalities and how they can be integrated in an SbD or in a hazard/risk assessment framework. Finally, a critical analysis of the identified and tested methods and tools was performed with the aim to derive meaningful conclusions that will contribute to the future development of *in silico* models for the assessment of hazards and risks of NMs.

### Methodology

This report was prepared using a step-by-step process to collect the relevant information and give EUON and its stakeholders the pertinent insights. For this reason, this project was divided into subtasks:

1. The structured and systematic literature mining, based on specific criteria of inclusion and exclusion of research studies defined in this project.
2. The critical analysis of the collected data, and the thorough analysis of the identified *in silico* tools and models with specific focus on the project requirements (e.g., data availability and their filtering, model validation and statistical evaluation, the definition of the models' domain of applicability, the performance of robustness tests, the mechanistic interpretation of the models (e.g., discussion on the causality between a property used within the model as independent variable and the toxicity of the studied NMs), the availability of standardised reports and the models' availability as user-friendly applications or web-services).
3. The identification of experts through the reviewed literature, research consortia, academic and industrial associations, and commercial networks to cover different stakeholder groups and the conduction of surveys, to collect state-of-the-art information on the available frameworks and on the scientific developments for read-across and other *in silico* approaches as alternatives to the testing of NMs.
4. The development of three case studies as examples for grouping, read-across, and *in silico* models, based on the previously identified studies, to illustrate the state-of-the-art and how these grouping, read-across, and/or *in silico* modelling can be used in SbD approaches or in NMs risk assessment.
5. The synthesis and presentation of the results, based on all acquired data from the previous subtasks. This report, therefore, summarises identified gaps, overall conclusions on current

---

<sup>a</sup>According to the Organisation for Economic Co-operation and Development (OECD) definition, "The SbD (Safe-by-Design, Safer-by-Design, or Safety-by-Design) concept refers to identifying the risks and uncertainties concerning humans and the environment at an early phase of the innovation process so as to minimize uncertainties, potential hazard(s) and/or exposure. The SbD approach addresses the safety of the material/product and associated processes through the whole life cycle: from the Research and Development (R&D) phase to production, use, recycling, and disposal." <sup>253</sup>

applicability and an outlook on how to drive future developments in the areas of read-across and alternative *in silico* approaches.

Through the methodology followed in this project we aim to ensure that the results from scientific literature and industry reports in the field are collected and analysed in a systematic manner and that they are presented in an informative transparent way. By using a comprehensive and standardised approach to literature mining these results can be considered as most up to date and relevant in the field. Thus, ECHA/EUON and its stakeholders are supplied with timely information on the existing *in silico* methods, and their applicability in NMs hazard and risk assessment applications.

## Results

The objective of this project was a comprehensive analysis of available computational models and *in silico* methods and tools for hazard and risk assessment of NMs through a systematic literature review, interviews with experts in the field and the development of relevant case studies.

In total 190 models/methodologies were assessed including nanoQSAR, grouping and read-across methods, AOPs, PBPK models and simulations. The models were retrieved from peer reviewed scientific publications and from the publicly available reports/websites of international programmes and research projects. Information was extracted from the publications and reports and was systematically collated in a spreadsheet file. Charts are provided -whenever possible- to better capture the key findings of the analysis.

Interviews with relevant experts were conducted via an online survey on the existing read-across and other computational methods for the risk assessment or SbD of NMs. The questionnaire of the survey was prepared in collaboration with ECHA/EUON based on the findings of the literature search and questions were targeted to address as many of the project's goals as possible. In total 36 experts participated in the survey. Their responses were collected, analysed, and compared with the results of the literature search.

The next paragraphs summarise the key findings of the abovementioned analysis, as well as suggesting potential actions in order to offer reliable, robust and high-quality models to serve the NMs hazard and risk assessment.

## Key findings

### Literature review

- In their majority (82%) the quality of the assessed models and of their reporting are quite high. When evaluating the key model aspects, such as the data pre-processing, the models' validation and interpretation, and the definition of their applicability domains, satisfactory results and conclusions were derived.
- Many efforts have been made in the field of nanoinformatics to develop methods that are specifically designed for NM datasets, taking into account the unique characteristics of the NMs (the properties of materials at the nanoscale, the NMs dynamic behaviour depending on the environment, the data scarcity). In addition, work has been done to develop methodologies and algorithms of general use that have been proven to produce reliable predictions even when small datasets are used. In fact, these methodologies can be employed to easily create different models based on different input datasets, allowing their use in different cases (endpoints, properties, or types of NMs) without any need of a grouping hypothesis.
- Automation and optimization of the *in silico* methodologies and modelling workflows have also been considered, to create models with improved features that converge to optimal grouping hypotheses and produce accurate predictions. Furthermore, the advances in artificial

intelligence (AI) and in machine learning have been integrated in the development of predictive *in silico* workflows. For example, deep learning approaches are already applied in the field of computational NM evaluation to extract information from microscopy images and later predict physicochemical parameters or possible NMs adverse effects.

- Progress has been also made in the development of other alternative methods such as simulations, PBPK models, and AOPs that unravel the NMs toxicity mechanisms and help to understand the NMs exposure routes in different organisms. Such methods greatly contribute to the SbD of novel NMs.
- It has been also proposed that workflows of mutually dependent models which are linked through specific properties/descriptors can be created to fill data gaps and reduce the need for experimental evaluation/or testing for already existing datasets (structure-activity prediction networks-SAPNets). These networks can be constructed from the endpoint to the descriptors that describe the NM structure and thus, later they can be integrated in NMs SbD processes as they can give enough details on what structural properties need to be changed to produce NMs with desired properties.
- The cornerstone for the development of a predictive model is the data availability, especially for nanoQSAR and read-across methodologies that are based on data-driven machine learning methods. In this course, input data of low quality (e.g., noisy, incomplete, free of bias, and inaccurate data) will inevitably lead to incorrect predictions thus, researchers seek to employ datasets that are curated and accepted by the scientific community. Based on the literature search results, we observed that even if simpler or more complex modelling methods and algorithms have been devised (e.g., incorporating the NMs particularities), due to the general nanomaterial data scarcity or datasets unbalance, they are developed and validated on a limited number of curated datasets found in literature. Therefore, their ability to estimate endpoints of interest in real-case applications is not widely assessed, and this hinders the models' regulatory acceptance.
- Apart from nanoQSARs and read-across models, for PBPK methods the data availability is crucial for their validation and consequently their regulatory acceptance because they are based on the comparison of modelled outcomes with conventionally retrieved (*in vivo* and *in vitro*) measurements. The limited data and consequently the limited applicability domains, may prevent the timely prediction of the NMs effects on the humans' health and safety. For example, this is the case of computational methodologies that assess the molecular mechanisms leading to adverse effects (AOPs), the interaction mechanisms between NMs and biomolecules (MD simulations) and the biodistribution of NMs in an organism after exposure (PBPK models): The data scarcity may prevent the timely identification of such mechanisms and consequently the timely prediction of the NMs toxicity.
- Again, in the development of SAPNets, meta-data scarcity prevents the development of predictive workflows consisting of already existing models, as their training data (the data used to build the models/algorithms on and adjust their hyperparameters) may be measured under different conditions that are unknown to the stakeholders. A characteristic example is presented in this report regarding the use of the NMs zeta potential as a linking property between models. The zeta potential is routinely measured during NMs characterisation and is linked with their stability in a liquid medium (e.g., it controls NMs tendency to form agglomerates) and in general it can regulate the NMs behaviour and functionality<sup>1</sup>. In this course, several models have been developed for the NM zeta potential prediction, that could be useful for NMs SbD. The zeta potential depends -among others- on the pH and the concentration of the suspension, the temperature, etc. These meta-data are usually not available to the modelers or the users. This renders the prediction of zeta potential deficient in terms of the conditions where the prediction refers to. It also makes it impossible to use as a linking property to combine models under the same predictive workflow. In such cases, it must be ensured that in the two combined models the zeta potential will be measured under the

same conditions, otherwise, meaningful predictions are not going to be produced. This applies for every linking property between models.

- To ensure the reproducibility of the *in silico* methods, all the necessary steps, parameters and assumptions/approximations should be reported for the model (ideally using a standardised template). In many cases, this reporting is missing: a typical example is the lack of documentation of the exact NM samples used during model training and validation especially when they are selected randomly (in a non-reproducible way) from a pool of available data.
- Different phases of the same material (different crystal forms) can lead to different extrinsic properties. These crystal forms are described through space groups that encode the arrangement of a material's atoms in space and should be provided to characterise the NM in study. In addition, in the case that dataset enrichment with molecular descriptors is sought, to encode into numerical values the NMs' structural characteristics, it is not possible to produce them if the NM space groups are not known. In many cases in the reviewed publications the space groups of the NMs are not available.
- The use of a standardised reporting template (e.g., QSAR Model Reporting Format-QMRF or the Modelling Data-MODA reports) could be a measure to overcome the reproducibility issues, however most of the assessed studies did not provide such a report.
- Model validation is a crucial step during the *in silico* investigation of NMs, as it evaluates their accuracy and suitability for use with real-case data and it contributes to the models' regulatory acceptance. In this course, the use of questionable performance metrics might not accurately represent how reliable a model is in practice. For instance, the F1-score which is popular for the validation of machine learning classification models and is used to validate many of the models collected in this review, has received criticism. In fact, it lacks "symmetry" and varies when the positive class is switched with the negative class in the confusion matrix (the values of true positive-negative and false positive-negative are also switched) and also may hold misleading information when the imbalanced datasets are treated. Its use as a statistical measure should be reconsidered or -if used- it should be introduced in a broader scheme of multi-fitness criteria.
- Since there is no one-size-fits-all performance metric for each and every modelling application, choosing one over others results in a number of compromises. It is advisable to present a variety of metrics (in this review we considered at least four) to obtain a complete and thorough validation of the prediction performance to properly assess the quality-of-fit and predictivity of a model. Many of the reported models did not include enough statistical measures.
- To avoid misinterpretations and make meaningful comparisons between models when presenting the statistical evaluation of a model, definitions and formulas should be provided (e.g., the use of the "R-squared" term should be avoided without providing additional information. The definition -coefficient of determination ( $R^2$ ) or squared Pearson's correlation coefficient ( $r^2$ )- and the respective formula should be also provided. This applies also for the formulas of other statistical metrics.
- The disseminated models (in contrast to the models that are not yet available via a user-friendly format) can maximise their utility and serve in the future for the design of sustainable and benign NMs. Most of the assessed models can be implemented as user-friendly tools considering that the relevant codes and scripts are available through public repositories, supplementary files, or equations in the scientific publications. In addition, 38 models and methodologies were identified to be already available through a user-friendly environment (e.g., as a web-application, a stand-alone tool, etc.). The 38 models were tested and were briefly presented within the report. The majority of these tools are simple to use by non-informatics experts as they use common user-interface (UI) elements. Stakeholders are more comfortable to use the tools through menus, buttons, radio-buttons, etc. and they can observe the results in tabular format or through automatically produced plots and graphs. The format

of the results permits their storage in the users' computers and their further post-processing; thus, it is easy to include them in a future NMs SbD, hazard or risk assessments.

- Nonetheless, the model's dissemination as a user-friendly tool should not be considered a secondary task when developing a predictive model. Some of the assessed tools present weaknesses during their use, such as inconsistent input, impractical user-interface, server timeouts, etc. These flaws may prevent stakeholders from using them.
- The deployment of an *in silico* model as a software component (e.g., web-application, stand-alone tool, etc.) requires expertise in the development of Graphical User Interfaces (GUIs), in the storing and retrieval of data and information, in server protocols, etc. Modelers and computational scientists may lack this kind of expertise, and this may not allow them to disseminate their models via a user-friendly environment. One of the positive steps taken in the field is development of deployment and hosting platforms such as Enalos Cloud Platform (<http://www.enaloscloud.novamechanics.com/>), Jaqpot (<https://www.jaqpot.org/>), and QsarDB (<https://qsardb.org/>). Through these platforms it is possible to deploy and share the developed nanoinformatics models and ensure their long term preservation and accessibility from the stakeholders. It is also possible to overcome issues with GUI impracticalities (as the ones mentioned before). Finally, these platforms contribute to making models **F**indable, **A**ccessible, **I**nteroperable, and **R**eusable-FAIR. Many of the identified models available as user-friendly tools are available through these platforms.
- The model's dissemination as a user-friendly application or tool is not sufficient without supplying training material on how to use the tool and interpret the results. Many of the presented tools in the appendix of this report did not provide any training material (e.g., video tutorial, written manual, etc.). In some of these cases a short presentation of the tool is included in the relevant scientific publication however, a more detailed manual could be very beneficial for the stakeholders, specifically on topics such as how to select the input parameters and how to interpret the results.
- As stated before, the majority of the analysed cases involved the definition of the applicability domain as a step in the regular modelling procedure. However, a remarkably large number of studies (44%) did not specify the models' applicability boundaries, raising doubts on the reliability of the generated predictions. Another issue in some of the tools is that, although the domain of applicability is defined within the model development workflow (e.g., included in the relevant publication or standardised report), it is not included in the tool, for example, through an indication of the reliability of the predictions based on the applicability domain limits. Without these limits it is not possible to increase the users' confidence in the produced results/predictions and this may prevent them from using the respective tools in real life applications.

### Experts' opinion

- Firstly, experts both from academia and the industry responded to the questionnaire, and the 69% of them claimed to have employed *in silico* methods or tools for research, for regulatory or commercial purposes, using a variety of tools. Also, the experts were involved in the development of nanoQSAR, grouping, and read-across models, as well as PBPK models, molecular simulations, mixture toxicity prediction models and (quantitative) AOPs, therefore a broad range of the methodologies was covered in this report. As for the development of the *in silico* methodologies, the majority of the experts seem to use the recommended OECD guidelines for validation purposes; both internal and external validation techniques were reported in the relevant questions. Likewise, most of them use several statistical measures to evaluate and ensure the credibility of their models, along with providing the applicability domain and a QMRF/MODA report. In general, model developers seem to be in line with the guidelines for the computational methods validation and dissemination.

- The NMs data scarcity and heterogeneity issue was also highlighted by the experts. In fact, in their majority they agreed that experimental NM data are limited, scattered across different sources and in different formats, and not collected or harmonised under a common protocol. The protocol of data collection and curation refers to a predefined procedural method that is followed for designing and conducting an experiment, and for analysing the results. It allows experimental replication and permits researchers to assess data consistency when they are collected from different sources.
- Meta-data (e.g., the experimental conditions under which descriptors are measured or the assumptions made for molecular descriptors calculations) are important for data organisation, consistency and searchability, but are often also unavailable. This makes it difficult for the stakeholders to fully understand the experimental system that is modelled, impedes data organisation in central repositories and also leads to the development of models for a limited number of NMs with limited applicability domains.
- The data scarcity could be ceased by data collation (combination of existing datasets in an extended set) and more reliable models or models with extended applicability domains could be developed. To achieve so, the included NM samples and/or descriptors in two datasets under consideration should be consistent and compatible, to ensure that the data represent the same information. However, the meta-data scarcity/unavailability prevents data collation to enrich existing datasets, e.g., following the example of the NMs zeta potential, if the experimental conditions under which it is measured are unknown, it is not possible to combine its values for different NMs under the same descriptor. Thus, the model development for larger and/or more diverse datasets is not possible.
- Despite of the weaknesses concerning NM-data availability, 70% of the responders reported that data quality control is a standard procedure prior to modelling, through various data curation and pre-processing steps. Another encouraging finding is that experts prefer both FAIR and Open Data to develop their models, which showcases a preferability towards transparency and data accessibility. Both cases demonstrate that model developers are careful about the quality of data and take steps to cope with data scarcity and known issues with data quality, in order to enhance the reliability of their models.
- As far as the dissemination of the developed *in silico* methods as user-friendly tools is concerned, nearly 80% of the respondents highlighted the necessity to support the regulatory applications and the SbD of novel NMs, therefore experts who developed such methodologies reported that they provide them, as web-tools and software. About three-quarters of experts agree that model dissemination as a user-friendly tool is necessary. More specifically, more than half of the responders reported that they include the source code for their developed tools, while 40% of them provide APIs for the remote accessing of their tools. Moreover, even though some modellers require a one-time-fee or a subscription for using their tools, the striking majority are willing to make their models freely available to stakeholders. Overall, experts demonstrate a mentality towards providing user-friendly workflows, web-tools, and software, aiming to increase confidence and facilitate accessibility for stakeholders.
- Nonetheless, experts also highlighted the reluctance of the stakeholders to use the available models and tools as they are not systematically validated or independently assessed, and they are not linked to real cases and applications.
- Interestingly, the majority of the experts themselves claimed that they are not aware of any *in silico* methodologies integrated in industrial, regulatory or research level for the SbD, grouping, and read-across of NMs.
- The experts' feedback can be summarised in the following opinion: "There is no open science core: no open data core, no open-source core, no open standards core". Specifically, the absence of community standards regarding Open data and models hinders the regulatory acceptance of the alternative *in silico* methodologies and their extensive use in NMs SbD and risk assessment.



## Case studies

- Three case studies were conducted for illustrating how the identified *in silico* methodologies can be employed in NM-related applications. In the first case study, we developed read-across models or used tools reported in the literature for the assessment of the anti-microbial activity of various carbon-based NPs. The interactions between carbon-based NPs and a SARS-CoV-2 RNA fragment were quantified for the examination of the potential of stabilising the cov-RNA fragment. The employed methodologies were the “EnaloskNN” integrated in KNIME platform, secondly, the Apellis web-tool, and a quantitative read-across method by the DTC Lab.
- We observed that all three methods in this case study produced quite similar results in terms of validation statistics, indicating that the performance of the models is consistent. Also, all three models are adequately validated, thus encouraging the stakeholders to trust the results and use the generated predictions. Consistent performance across the models is also observed from the domain applicability, since the two models that offer a reliability indication based on their applicability domain, EnaloskNN and Apellis, seem to provide 100% reliability for the testing set, implying that similar NMs’ results are likely to be reliable. Additionally, those web tools deploy interactive visualisations for the interpretation of results in a user-friendly environment. By bringing together the predictions from two or more different models in a consensus framework, it may reduce the bias of each individual model. Therefore, the assessment of the endpoint is more trustworthy, and the stakeholders can have greater confidence in the final results.
- The second case study focused on the assessment of six different tools available as ready-to-use online applications for the prediction of TiO<sub>2</sub>-based NMs’ endpoints. The availability of these computational models through a GUI permits non-informatics specialists to use them to design novel NMs and for risk evaluation without involving any complicated processes and technical details. Thus, the models are addressed to a larger audience, both for researchers and designers, and they allow faster sensitivity analyses or virtual screenings of newly designed NMs. By providing a practical and comprehensive interface, those tools facilitate predictions of various endpoint and toxicity-related properties, and they enable quick decision-making by the stakeholders. Additionally, the different TiO<sub>2</sub>-based are suitable for use in a safe-by-design framework, since they allow the exploration of the correlations between different properties or the prediction of the toxicity of a largely used NM.
- On the other hand, after the assessment of those computational tools, a few drawbacks were identified which may discourage the stakeholders from using them. For instance, some tools did not provide the limits of the applicability domain or had inconsistencies with the original publications of the studies. In addition, most of these tools had inadequate tutorials, that would not help/guide the stakeholders to use them properly or to make use of all the available functionalities (e.g., the tutorials include too many technical details for the tool, or they lack a proper results interpretation). In other cases, no manual to use the tools is even offered.
- In the third case study, we developed a predictive workflow based on the SAPNets scheme of interdependent predictive models. This workflow combines two or more existing nanoinformatics models in such a way that an input descriptor of the final model is predicted by another meta-model. SAPNets facilitate the utilisation of available models from the literature and contribute to the nano-data gap-filling since no experimental evaluation is needed. A further advantage of SAPNets, as seen from the case study, is the fact that it allows a more comprehensive analysis for the factors that influence other NM properties since they consider a wider range of descriptors. They also offer better guidance for safer NM design, provided that they highlight the structural characteristics, such as the size and the coating that affect the desired property, thus novel NMs with specific attributes can be designed with minimal error.
- However, the applicability domains of the combined models should be clearly stated in the individual models to ensure the reliability of the models, and error propagation should be

calculated across the SAPNet. Additionally, details of the experimental conditions that regulate NM behaviour should be provided to ensure the compatibility of the combined models. Experimental and computational data and meta-data are essential for incorporating two or more separate *in silico* methods.

## General conclusions

- The assessed studies in their majority are in a satisfactory quality level and efforts have been made to integrate in the developed computational workflows optimisation functions, and advanced methodologies (e.g., deep learning).
- Work has been done in the models' dissemination as user-friendly tools, and this may contribute to their use from stakeholders in regulatory and industrial sectors. The availability of models' scripts in public repositories in combination with the existence of deployment platforms can also accelerate their integration in different research activities and expand their use from interested users. Nonetheless, there is also room for improvement regarding the inclusion of information on the applicability domain of the models within the tools and the availability of training material for each tool.

Both the literature analysis and the experts' opinion converged in the following challenges:

- Data barriers exist regarding the available nanotoxicity-related datasets. Compared to conventional chemicals for which database solutions already exist (e.g., PubChem), standardised, structured, annotated, and reliable nanomaterial-related data are limited, and they are not always accessible. Data are found in different formats and there are consistency deficiencies between the data presented in different sources (e.g., data for the same endpoint are collected in different experimental conditions, under different protocols, or meta-data are not available). This prevents proper model development and validation, as well as datasets collation and enrichment.
- Apart from the data-related issues, the lack of models' reporting -including all the relevant information for model development, validation, and possible assumptions- poses barriers in their proper use from stakeholders.
- Setting aside the aforementioned issues, one of the greatest challenges in the field of the NMs *in silico* assessment, is their integration in the industrial, regulatory or research activities for the SbD, grouping, and read-across of NMs. The confidence of the relevant stakeholders in the developed models is still low, mainly due to the limited modelling information from the developers, the lack of sufficient data to provide a good applicability domain, and the lack of validation with conventional experimental methods.

## Outlook

From both the literature research and the interviews, it was derived that the data-related limitations, as well as the lack of proper model reporting, create several issues (limited validation possibilities, reduced applicability domains, difficulties in datasets enrichment and models' combination) that impede a wider regulatory acceptance of the computational methods for NMs risk assessment. For this reason, the following suggestions are intended to assist in the process of the regulatory approval of the models.

- The key to resolve the identified flaws is transparency at all levels (for generated data and models). This may include the accessibility of the datasets enabling the fast computational assessment of several NM categories, expanding the number of the already *in silico* assessed NMs. Existing nano-databases (e.g., eNanoMapper, NanoPharos, nanoHUB, etc.) can be employed to host *ready-to-model* datasets. In fact, some of the models assessed in this project, were built using data that are available in some of these databases.

- Transparency can be further improved by making data and meta-data FAIR. Finding data is the first step towards reusing it. Both humans and computers should have no difficulty finding data and meta-data. This step in the FAIRification process is crucial since machine-readable meta-data are required for the automatic discovery of datasets and services. Once the user has located the necessary data, they must be aware of how to access it, which may involve authentication and authorization. Usually, the data must be combined with additional data and must work with workflows or apps for analysis, storing, and processing, thus knowledge should be represented using a “vocabulary” that permits interoperability. FAIR’s main objective is to maximize data reuse. To accomplish this and support future combination of datasets, meta-data and data must be adequately described to be replicated and/or integrated across multiple settings. Many repositories may be used for FAIR NM data storage (e.g., eNanoMapper, caNanoLab, and NanoCommons).
- The support of the NInChI<sup>2</sup> initiative (a machine-readable identifier to represent a particular group of NMs considering its necessary characterisation properties based on the International Chemical Identifier (INChI) concept) could also contribute to the generation and distribution of FAIR nano-data.
- The development of novel nanodescriptors that encode and integrate the different NM components (chemical and structural information, interactions with biomolecules, etc.), as well as the development of computational methodologies that consider such descriptors will support the reliability and interpretability of the models.
- The modelling workflow (including data pre-processing or filtering, data splitting, validation, any robustness or sensitivity analysis tests, domain of applicability definition) should be clearly presented either via a scientific publication or through a report that accompanies the model.
- In addition, when all the modelling information is presented in a transparent and understandable manner, stakeholders are able to understand easily if two or more models are suitable to be combined in predictive workflows before committing too much time and effort in the process of studying all the different aspects of the modelling procedure or searching for the training data.
- Considering the current situation, it is necessary that researchers are particularly cautious and that they study thoroughly the respective parameters when combining existing models into predictive workflows to ensure their compatibility and the generation of reliable predictions across the workflow (e.g., compatibility of linking descriptors, error propagation, applicability domains).
- Modellers need to provide as many statistical measures as possible, in order to map the actual performance of the models, including the weaknesses of the selected approach.
- The communication between different stakeholders is also important. Modellers are often not specialists in the field of nanotoxicity, thus a communication and collaboration channel between modellers and specialists will contribute to the interpretation of the toxicity mechanisms revealed by the models, or the mapping of the descriptors’ space.
- Communication between modellers and specialists will also contribute to the improvement of the developed tools. Feedback can be given on the different tool features and/or other suggested applications may be developed, to expand the usefulness of the tools in real case scenarios.
- Stakeholders of the models and tools usually have different backgrounds from the computational scientists. Thus, the use of the developed models/tools cannot be taken for granted. Effort should be spent in the preparation of training material that accompanies each tool, where its use is presented in detail and in simple language. The necessary scientific or technical terms, the input, and the expected output can also be provided through the respective manual. In cases where the scientific group developing a predictive model lacks time, resources, or know-how to build a tool with a graphical user interface, they are strongly encouraged to share their code in public repositories (such as GitHub), so that other teams can

access, modify -if necessary- and disseminate the models as user friendly applications (e.g., as previously stated, QsarDB, Jaqpot and Enalos Cloud are deployment platforms where models can be uploaded as user-friendly web applications).

- To assess the level of readiness of the developed tools in terms of their regulatory acceptance, the Transparency, Reliability, Accessibility, Applicability and Completeness-TRAAC framework can be employed. This framework evaluates the tools through a series of criteria that quantify their compatibility with regulatory frameworks and their usefulness and usability for end-users. It also identifies, through its scoring system the possible limitations that impede the tools' regulatory acceptance, contributing to their future improvement<sup>3</sup>.
- To enhance the trust on the computational results of the different user-friendly tools, it is important that they are frequently updated with newer input or data. Outdated tools are possible to be rejected by the stakeholders.
- Using the *in silico* tools in real-life cases or in comparison to results from conventional/empirical test methods (e.g., *in vivo*, *in vitro*) is another technique to support regulatory acceptance and finally encourage their employment as NM testing supporting methods and/or alternatives. The necessary communication to perform such testing can be established by taking advantage of the existing network of computational teams and laboratories with collaborators or consortium members in the Industry, Academia, or Regulatory Agencies. We should note here that this can be a short-term procedure that will not increase the use of laboratory animals. In addition, it is not necessary to perform *in vivo* experiments from scratch: industries and laboratories that perform NMs risk assessment on a regular basis, can employ the *in silico* tools that are already available, and compare the predictions to their previous experimental results to enhance their trust in the methods. A campaign from national and international institutions highlighting (possibly through success stories involving the design/optimisation of NMs based on computational methodologies) the time and resources that could be saved from the use of computational methods as alternatives for NMs risk assessment, could spread their use to the stakeholders.
- Frequently, lack of time or a scientific journal's publishing guidelines do not allow the extensive and detailed presentation of all the model aspects. For this reason, the use of standardized reports covering the key model aspects should be a prerequisite when a predictive model is submitted to be published. In addition, the use of a standardised format when reporting a model allows the organisation of the necessary information in a straightforward manner.
- While reviewing some of the models and tools presented in various publications, we identified certain flaws and errors. We recognize that the peer review process for scientific articles can be challenging and that it is possible for basic (e.g., inadequate model validation) or minor errors (e.g., model parameters typos) to slip through. Therefore, it is important to improve the reviewing system so that the highest standards of scientific rigor can be upheld to maintain the integrity and quality of scientific research (e.g., through seminars for academic writing for authors, formal reviewers training, reviewers rewarding, use of AI tools to detect artificially generated text or manipulated plots and images before the manuscript is sent to the reviewer, etc.).
- Finally, to ensure that the requirements are met (e.g., reporting of all modelling aspects including data used, assumptions, validation, evaluation metrics, applicability domain limits, code availability), peer-reviewers should be encouraged to specifically check if the above-mentioned criteria are met by the author. In case they are not included in the publication, the reviewers should encourage the authors to supplement their work with this key information to make it more meaningful, applicable outside research and to contribute to the further development of computational methods to support the domain of nanosafety. Publishers should also update their reporting standards. For example, *Nature* requires that the authors provide their code and data through public repositories to improve transparency and reproducibility of the results<sup>4</sup>.

## 1 Introduction

The rise of nanotechnology applications has led to the discovery of numerous novel materials that are different from their bulk counterparts with respect to their physicochemical properties. This has resulted in a rapidly expanding spectrum of commercial uses that take advantage of these novel properties (such as cosmetics, food additives, solar cells, paints, etc.). Nonetheless, the extended use of NMs in everyday life products could have consequences if a connection with long-term health and environmental hazards is found, or even simply suspected. Given the expense and time constraints involved with the experimental hazard and risk assessment, which commonly involves using animals, many NMs have not yet had their potential harmful biological effects thoroughly evaluated. The 3R (**R**eplacement, **R**eduction, and **R**efinement) principles were created as a framework for conducting more ethical animal research<sup>5</sup>. In this course, animal replacement can be achieved by speeding up the creation and application of reliable, accurate computational models and tools that are based on the most recent research and technological advancements. Therefore, predictive nanoinformatics modelling can greatly contribute to the replacement of animals in NMs hazard and risk assessment.

The main purpose of this study is to inform on current frameworks and current developments for read-across and other computational methodologies, which are used as alternatives in NMs hazard and risk assessment, and in the NMs safety-by-design (SbD). The models and tools that are currently available for the *in silico* assessment of NMs were identified, evaluated, and discussed in this regard through the use of a structured literature study.

The findings presented in this report are based on a carefully balanced combination of primary and secondary research for tools focused on the computational investigation of NMs, as well as on a hands-on testing of the identified nanoinformatics tools or methods through the development of case studies. In the primary search, we directly contacted the relevant experts from Academia and Industry who are involved in the development of computational methods or are using such methods for the *in silico* investigation of NMs. We collected their opinion through a carefully prepared online questionnaire. In the secondary search, we conducted a thorough literature review to map the existing methodologies and tools that have been developed in the field of nanoinformatics, focusing on specific aspects of these models or modelling methodologies, such as:

- The model development and validation process,
- The existence of an applicability domain,
- The availability of modelling data and the dissemination of the models as user-friendly tools,
- The availability of information to reproduce the models and,
- The models' possible data gaps and weaknesses, etc.

Based on the findings of the primary and secondary search, three case studies were developed with the aim to assess a representative number of nanoinformatics tools, to test their functionalities and generated results, and to discuss how these tools can be exploited in the NMs SbD and risk assessment.

The results obtained for the literature research and for the case studies were built to provide a robust and validated analysis of the current state of the existing (read-across) models and tools for the *in silico* investigation of NMs. We particularly focused on the existing gaps and on the weaknesses that could be improved in the future, to further enhance the confidence of the stakeholders and their actual use on NMs risk assessment processes.

This report is structured as follows:

- First the main types of computational methodologies and their different characteristics are presented to the reader with their respective assessment.
- Next, the literature review (secondary research) methodology is provided, as well as the information extraction, collation, and its systematic indexing process.
- The analysis of the collected publications and tools is presented next for each family of analysed models, focussing on specific aspects of the models. Especially for the nanoQSAR and grouping/read-across models, emphasis has been put on data and their availability/accessibility, the modelling methodologies, the model validation, the domain of applicability (DoA) and dissemination to stakeholders. Subsequently nanoQSAR and read-across models are put through a quality assessment according to the above-mentioned criteria.
- The findings of the primary research (interviews with the experts) are then provided with their respective analysis and outcomes with regard to the choice of the case studies.
- Finally, the process of developing and the results of the three case studies, where a substantial number of methods and tools is tested, are presented and discussed.
- Based on the above results, conclusions are derived and presented in detail.
- Finally, a list of Appendices (I-VI) is added to the report to supplement the work, including information on the assessed studies and the interviews, extra graphs and tables that may be of use for the readers.

## 2 Main types of assessed *in silico* methodologies

Even though NMs are extensively used in a plethora of products because of their special properties, several indications have been detected linking them to environmental, health and safety hazards. Therefore, many *in silico* methods have been developed to predict the risks for humans and the environment, posed by NMs, as an alternative to *in vivo* and *in vitro* animal testing. In this section we are presenting the results of the analysis of the computational models and methods collected during the literature mining. A brief presentation of the included families of models is presented here before proceeding to the presentation of the results of this project.

### 2.1 nanoQSAR

A quantitative structure–activity/toxicity/property relationship (QSAR/QSTR/QSPR), is a field of molecular modelling that employs statistical methods and machine learning to explore the connections and establish a mathematical function or a more complex model between chemical structure and biological activity<sup>6,7</sup> or another property. The chemical structure of the studied molecules is encoded mathematically in the molecular descriptors and the activity is often known as the “endpoint”. QSARs have been widely used in the fields of drug discovery and chemical toxicity. In the area of nanotoxicity the concept of QSARs (known as nanoQSAR or quantitative nanostructure-activity relationship-QNAR) has been adapted to the specificities of NMs aiming to correlate several NM properties (physicochemical properties, theoretical molecular descriptors and structural descriptors of the NMs core and coating) to observed toxicological effects or other properties<sup>8,9</sup>. The output result of nanoQSAR models can be either qualitative, by classifying a NM e.g., as “toxic” or “non-toxic”, or quantitative, by predicting a numerical endpoint value. After obtaining the nanoQSAR model, it can be applied to a NM which had not been previously tested or was not included during the model development. Critical points a nanoQSAR study must contain, is the model validation, namely internal and external predictive ability, and the modelling approach (e.g., linear or non-linear regression). When analysing the results of this review we clustered all the QSAR-type models (QSTR, QNAR, QSPR machine learning, etc.) developed for NMs under the category of “nanoQSARs” as we considered that all these abbreviations are actually expressing the same family of models.

### 2.2 Grouping and read-across

The establishment of nanoQSARs is not a universal solution in the area of *in silico* nanotoxicity assessment. Due to a scarcity of experimental data, only a few theoretical descriptors are known for NMs. Thus, the proper application of nanoQSAR statistical techniques, including the essential external validation of the model, is frequently hindered by a lack of data<sup>6,8</sup>. Small data sets tend to produce models that are more likely to be overfit and have narrow applicability domains, which makes them unsuitable for estimating novel NM parameters. In addition, the heterogeneity of the different NM structures prevents the development of universal nanoQSAR models<sup>8</sup>. Grouping and read-across approaches emerged as alternative methodologies for data gap filling especially in cases of limited datasets. Read-across models are semi-manual approaches for making data-based predictions<sup>8</sup> and are based on the concept that by interpolating or extrapolating the property of one or a group of NMs (analogue or source substances), the value of a similar NM (target substance) can be obtained with satisfactory accuracy<sup>10</sup>. Similarity between NMs is quantified based on the available physicochemical, atomistic, molecular, periodic table properties, image analysis descriptors, etc. The Read-Across Assessment Framework (RAAF)<sup>11</sup>, created by the European Chemicals Agency (ECHA), is an internal tool for evaluating predictions of substance attributes based on read-across in the context of the Registration, Evaluation, Authorization, and Restriction of Chemicals (REACH) Regulation. To harmonize the emerging read-across techniques, ECHA also released a document including a systematic workflow for NMs grouping and read-across, consisting of seven steps (Figure 1)<sup>12</sup>.

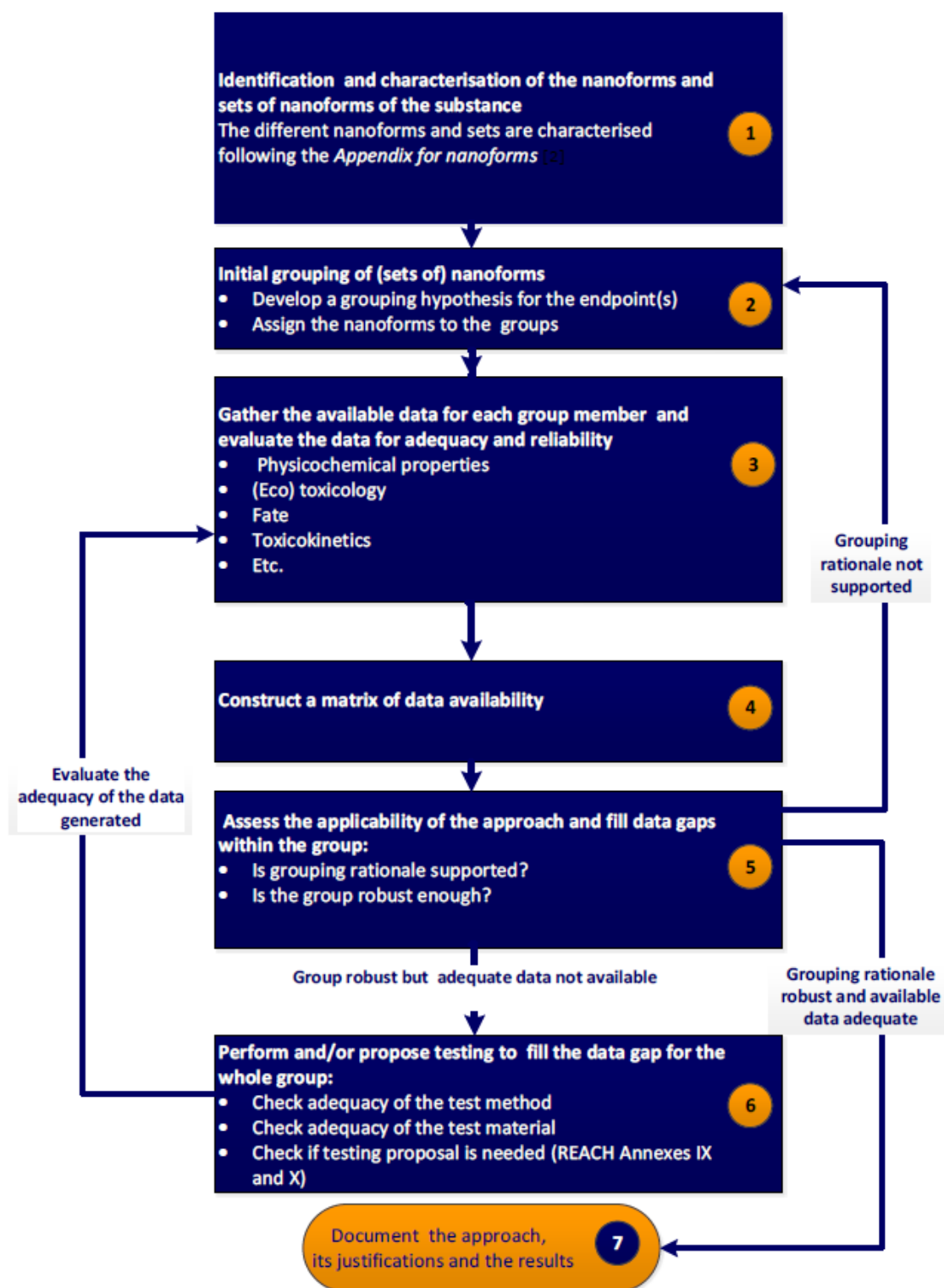


Figure 1: ECHA's recommended stepwise approach for the systematic and transparent data collection and documentation of read-across approaches. Image derived from the "Appendix R.6-1 for nanoforms applicable to the Guidance on QSARs and Grouping of Chemicals"<sup>12</sup>.

### 2.3 Physiologically-based pharmacokinetic modelling

Another approach for predicting toxicokinetics and potential hazard of NMs is the physiologically-based pharmacokinetic (PBPK) modelling. These models are based on the anatomical and physiological structure of the living system and offer an effective method for extrapolating dose and species as well



as for predicting biodistribution in the body over time<sup>13</sup>. PBPK models are essentially equations in different body compartments to evaluate how nanoparticles (NPs) behave after entering a living organism (humans and animals). In fact, each compartment represents a specific area of the body, such as a single organ or tissue (such as the liver), a group of organs or tissues (e.g., slowly perfused tissue), or a section of an organ or tissue (e.g., the intracellular space)<sup>14</sup>. To estimate concentrations of NPs in organs and tissues, it is considered that body compartments are connected with each other via blood flow. Later, according to the law of mass conservation for each organ/tissue, algebraic/differential equations are acquired and mathematically explain NP transport within the body<sup>8</sup>. The NPs clearance phenomena (NPs removal from the body) are also considered in PBPK modelling to simulate long-term or cumulative exposures<sup>15</sup>. Certainly, several physiological and biochemical parameters for each organ (such as regional blood flows, tissue volumes, experiment-specific parameters) are required in order to solve the final system of equations. Eventually, PBPK models facilitate the understanding of mechanisms and processes such as absorption, (bio)distribution, metabolism, and excretion (ADME)<sup>14</sup> in a biological system, as well as providing indications for toxicity caused by NMs. Due to the different ADME behaviour of small molecules (drugs) and NPs when developing PBPKs for nano, several modifications to the “classical” PBPKs should be considered (e.g., differences in the transportation mechanism and kinetics, in the NMs metabolism and the effects of NMs in the lymphatic system)<sup>16</sup>. For example, NPs may or may not dissolve into the biological environment thus different cellular uptake and intracellular kinetics apply for each case<sup>17</sup>. This is the case of Ag NPs that partially dissolve into the gastric fluid<sup>18</sup> resulting to a mix of Ag particles and ions that can be absorbed by the stomach and the intestines. These differences should be converted into modifications of the relevant PBPK model<sup>19</sup>.

## 2.4 Molecular dynamics simulations

The molecular dynamics (MD) simulations are a computational area for molecular modelling, used to simulate the interactions between NMs and cell membranes, proteins, and DNA by applying the principles of classical mechanics<sup>20</sup>. The basis of MD simulations is the solution of Newton's equations of motion for all the atoms that compose a system while bound to volume and temperature constraints. In fact, a “virtual experiment” is performed and by having visual evidence of the interactions between NMs and the biological environment, it is possible to reveal all the molecular mechanisms that take place under the studied conditions. Thus, it is possible to understand in an intuitive way the causes of cytotoxicity and to extract data such as the energy changes or binding free energies of NMs. In order to reduce the large system to a simplified unit cell, periodic boundary conditions are often used in the simulations. A force field that describes the system as a function of atomic coordinates is required to run an MD simulation. According to the type of force field used, there are three different MD strategies: coarse-grained (CG) simulations, where a simplified representation of the system is employed (groups of atoms are replaced by “pseudo-atoms”) to speed up computations, all-atom simulations, where forces are calculated based on intermolecular potentials, and *ab initio* simulations, where the forces between atoms are calculated based on first principles. An MD simulation is performed using special software packages, both commercial and open-source, (e.g., GROMACS, NAMD, and LAMMPS)<sup>21,22</sup>.

## 2.5 Adverse outcome pathways

The adverse outcome pathway (AOP) framework portrays a series of causally related key events (KEs), between a molecular initiating event (MIE) and an adverse outcome (AO)<sup>23</sup>. A MIE is a special type of KE that explains the initial interaction between a stressor (typically a chemical compound/NM) and a biological target within an organism, one that causes perturbation and is the beginning of an AOP. The KEs are measurable and discrete changes in a physiological/biological state responsible for each step of the pathway. Relationships between two key events (KER) are used to connect one key event to another, by defining a causal and sequential relationship between them, and by determining the

upstream (causal) and the downstream (responding) event<sup>24</sup>. Finally, the AO can be defined at different levels: for human health effects, an AO rarely relates to whole population, but rather to specific organ damage (for example, liver fibrosis), which has consequences on the individual, whereas in environmental toxicology, the AO typically relates to growth inhibition, reduced survival, or reproductive impairment of an individual (for example, a fish), and the implications on the whole population<sup>23</sup>. The developed and assessed AOPs are systematically organised and shared through the AOP Knowledge Base (<https://aopkb.oecd.org/index.html>) and the AOP-Wiki (<https://aopwiki.org/>) which are publicly accessible<sup>25</sup>. It is important that an AOP framework contains the Event IDs of key events and AOs, which are assigned a unique number automatically upon creation and are documented in AOP-Wiki. Additionally, the weight of evidence (WoE) must be clearly assessed, by providing scientific support and explaining the degree of confidence. Lastly, the DoA must be addressed in terms of taxa, life stage, sex, the level of biological organisation and other aspects of biological context according to the Organisation for Economic Co-operation and Development (OECD) guidelines for AOPs<sup>25</sup>.

### 3 Extended literature search

To perform the literature search we developed a specific protocol with standardised processes. The protocol covered all steps including identifying, selecting, and critically appraising relevant resources, and synthesising the current body of knowledge on nano-specific *in silico* and read-across methods.

Structuring the methodology based on the specific protocol, ensured that:

1. The entire study is based on sound scientific pillars, which will enhance the reproducibility of the study,
2. The protocol will increase the credibility and transparency of the outputs of the review, so that they can be considered as the most relevant findings from scientific and industry literature in the field,
3. The study will provide access to timely information on the state of the art, and
4. The identified results are analysed, explored, and presented in a systematic, comprehensive, and informative way and visualised, where appropriate, with tables, graphs/charts, and datasets and, where possible, with infographics that are useful and interpretable.

A draft protocol was developed for conducting the review, by defining the relevant questions and scope, the methods of the review, and the eligibility criteria for the inclusion of studies/reports and materials into the study. This helped to reduce bias and the reviewers were committed to following the documented process. In addition, the extensiveness and reproducibility of the search strategy and the transparent reporting of how studies were selected and included in the study reduces bias in the selection of research studies. The search strategy is described in such a way that readers can estimate how much of the relevant literature is likely to have been found.

The review includes an assessment of the quality of the evidence in terms of study methodological soundness, which gives an indication of the strength of evidence provided by the review and allows emphasis to be given to the results from studies/sources of higher quality. The methodology of the review process is adequately documented to allow others to critically appraise the judgments made in study selection and the collection, analysis, and interpretation of the results and, if necessary, to repeat or update the systematic literature review.

Overall, the fundamental principles of the systematic reviews are closely followed, i.e.,

1. Methodological rigour and coherence in the retrieval and selection of studies/sources, assessment of their methodological quality, and the synthesis and interpretation of information,

2. Transparency, and
3. Reproducibility, as shown in Table 1.

Table 1: Characteristics and description of the review process.

#	Characteristics	Description
1	Study questions	Focused and explicit
2	Eligibility criteria for inclusion or exclusion of studies	Pre-defined and documented; objectively applied
3	Description of the review method	Reported and predefined in a protocol
4	Literature search	Structured to identify as many relevant studies as possible
5	Methodological quality assessment of included studies	Included, typically using a quality assessment tool
6	Reporting of study results	Full reporting of relevant results (numerical results)
7	Synthesis	Quantitative synthesis (meta-analysis) when possible

### 3.1.1 Search strategy protocol

The development of the search strategy protocol was based on the project's main objective, which is to provide EUON with insight on available frameworks and state of the art (scientific) developments for read-across and *in silico* approaches as alternative approaches to the testing of NMs. It also includes the development of these alternative approaches for exposure as well as hazard and risk assessment of NMs, through structured literature reviews. The study examined the available frameworks and read-across/*in silico* approaches already developed, as well as those under development (for which resources are publicly available). Based on the key aims of the project an initial set of key questions were identified that needed to be addressed through the review. These questions were used to define specific keywords to perform the literature search and evaluate the retrieved results, and which are highlighted in bold:

1. Which **frameworks** for **grouping** and **read-across** do exist already?
2. For these frameworks how were they **developed** and **validated**?
3. How far did this go and for which **purposes** or **application**?
4. What data or **physicochemical properties** and other ***in vitro*** or ***in vivo* properties** were used and were tested to build the **testing strategy**?
5. Were they evaluated against critical parameters for **reliability** and **applicability**?
6. Were any other properties (e.g., **atomistic**, **molecular**, **periodic table**) used for model development?
7. Were there **quality control** processes implemented regarding the data and **meta-data** used?
8. Were the data used or produced using **regulatory compliant** (e.g., OECD, ISO) guidelines?
9. Are the data used compliant with the **FAIR** (Findable, Accessible, Interoperable, Reusable) and **Open Data guidelines** of the EC?
10. Were the **applicability domains** assessed, communicated, and clearly explained?
11. Have the *in silico* approaches been documented using a specific template (e.g., **QMRF**, **MODA**)?

Based on these, the key search terms and the inclusion/exclusion criteria for study eligibility was defined to guide the entire process. Using a clearly defined protocol, which includes all the required search and evaluation criteria, we reduced bias, as the process was clearly specified in advance and the

reviewers were committed to follow it. This is strengthened further with the use of two independent reviewers for each resource. In case of disagreement, a third more senior reviewer was employed to provide the final decision on resource eligibility. It is noted that the reviewing team is multidisciplinary (chemists, physicists, pharmacists, bioinformaticians, chemical engineers), guaranteeing a thorough review. In addition, we reduced any bias in the selection of research studies by the extensiveness and reproducibility of the search strategy and the transparent reporting of how studies were selected and included in the project as presented in the section below.

### 3.1.2 Selection of databases to be examined

The first step in the review process was to identify search engines and DBs offering high-quality inventories on available frameworks and state of the art (scientific) developments for read-across and other *in silico* approaches. Indicative databases -in terms of access to academic journals, conference proceedings, etc.- or search engines that were mined include PubMed, Google Scholar, Scopus, UniChem, Scifinder, Reaxys, Web of Science, etc.

### 3.1.3 Definition of search terms and Boolean operators

Based on the project requirements (see key questions), we defined the relevant keywords to be used for retrieving the relevant literature. The search included a series of general searches, e.g., framework AND read-across AND *in silico* AND hazard AND risk AND nanomaterials. Further refinement included the defined key terms based on the project's objectives, i.e., the available alternative approaches for exposure as well as hazard and risk assessment of NMs. In this way, it was possible to identify and collect data on all project aspects. Furthermore, extensive literature search was performed for any identified challenges/gaps during the review procedure. This ensures that any missed literature with respect to the gaps/challenges was identified and reported, along with an estimation of the work performed in the field. This is particularly important as the review covered a substantial amount of time and there is a high chance that the identified gaps/challenges reported in literature would, in some cases, be researched already.

To maximise high quality data retrieval, we needed to identify the correct balance between the specificity and sensitivity of the research. These are key to an effective search strategy, with both presenting relevant advantages and disadvantages. A specific search provides a substantial amount of relevant research, while avoiding, to a large degree, irrelevant results. In this way there is a substantial amount of time saving when it comes to filtering and screening the results. The disadvantage is that the more specific the search becomes, the higher the risk of missing relevant literature. This is because specific search relies on searching very precise concepts and ideas and their combinations, while focussing on specific parts of the study, e.g., title and abstract. This can lead to data loss in the case that non-standard, modified, or novel terminology is used, or where a vague title and/or generalised abstract are present.

On the other hand, sensitivity provides researchers with the opportunity to capture most of the relevant literature and substantially lowers the risk of relevant data loss. The disadvantage, in this case, is the fact that more irrelevant literature is retrieved as well. As a result, the time and effort required for filtering and screening increases. This is because a sensitivity-focussed search relies on using more generalised terms and their combination and does not focus on specific study parts. This leads to more hits, both relevant and irrelevant, that need to be carefully screened to discard irrelevant studies. Summarising, a specific search is aimed at answering research questions with a high degree of certainty, using clearly defined and specific search terms. Sensitivity is required when the researchers are looking to perform an exhaustive literature search and when the concepts and the questions needed to be

addressed are not clearly defined. In our case, we need to find the optimum balance between the two concepts, while not restricting the search to one or limited databases or specific parts of literature.

### 3.1.4 Resources' selection criteria

As per the technical specification requirements, a clear set of resource inclusion and exclusion criteria has been defined to ensure high quality outcomes. These are based on the presented key questions, the project's requirements, and complemented with general search criteria. The current set of inclusion and exclusion criteria proposed (based on applicability) are:

Criteria for resource inclusion:

- English Language publications.
- Resources post 2010.
- Resources (including reviews) discussing alternative methods for hazard assessment of NMs and/or QSARs, chemical categories, grouping, read-across, Physiologically Based Pharmacokinetics (PBPK)/Toxicokinetics (PBTk) modelling, *in vitro* and *ex vivo* experimental results, AOP framework- alternative mechanism-focused methods in risk assessment.
- Resources from company websites, publicly available reports from research projects and full reports from the following H2020 and HORIZON projects: NanoSolveIT, NanoCommons, RiskGONE, Sabydoma, Nanogentools, CompSafeNano, Diagonal, WorldFAIR, NanoMILE, NanoFASE, etc.
- Resources providing information on the development and/or validation of alternative methods (including AOPs).
- Resources providing information on the boundaries, challenges, and/or gaps of relevant alternative methods.

Criteria for resource exclusion:

- Resources not in English.
- Resources prior to 2010.
- Resources in predatory journals (even if peer-reviewed).
- Non-peer-reviewed resources without references.
- Resources without full text access.
- Resources not describing the methodology used and required meta-data in full (research studies).

As stated before, both the screening and examination processes were evaluated, revised, and followed up by at least two individual reviewers, to prevent introduction of errors and personal biases, and if needed from a third senior reviewer. To facilitate the process the reasoning of rejection was documented by each reviewer, according to the given criteria. This ensures that the review was unbiased, and, in the case of disagreement or discovery of an error, it was easy to correct by other reviewers. While most duplicate findings were expected to be automatically removed, a minor number of duplicates remained and was identified during the evaluation and removed. The final list of references retrieved from the databases in total, the number of studies excluded in each review step, and the list of relevant references selected is provided.

### 3.1.5 Synthesis, presentation, and interpretation of data

Following the analysis of the collected studies based on the reported criteria, the research team proceeded to data synthesis, presentation, and interpretation. In the case of quantitative data, meta-analysis approaches were employed for analysis. When possible, the results were presented in tabular form using summary tables and visual tools (e.g., graphs, infographics) to assist with their interpretation and to ensure transparency of the process and its results. The research team was continuously

performing a well-structured analysis and discussion of collected data, to produce conclusions and their clear presentation, as part of the last step of the project.

## 4 Information extraction

After searching and identifying the research studies that would be included or excluded in the project, their thorough analysis was performed. Through an evaluated screening and examination process, we conducted a systematic extraction of information from the scientific publications or the publicly available EU and nationally funded project deliverables/websites. Specific aspects were identified based on which each reviewed study/tool was evaluated. The reported aspects when extracting the details from each source are the:

- Information on the read-across and *in silico* validated alternative methods (including type of modelling methodology, endpoint information, modelling data information on the dataset's size and descriptors).
- Information on data filtering (e.g., statistical or manual following experts' opinion) and variable selection.
- Information on the validation strategy (e.g., OECD guidelines) of the reported alternative methods.
- Reported statistical measures to evaluate the models' performance.
- Information on the alternative method applicability domain.
- Information on the applied sensitivity analysis or robustness tests, if any.
- Information on the mechanistic interpretation of the model.
- Information on the development and reproducibility of the alternative method, through standardised reports e.g., QMRF, MODA.
- Information on the availability of details to enable evaluation of the validity and suitability of the selected test methods.
- Information on the availability of the methodologies through a Graphical User Interface (GUI) (e.g., as web services) or as scripts in public repositories.
- Models' possible data gaps and weaknesses.

Each publication was examined in detail and evaluated by at least two reviewers in order to eliminate biases and errors. The assessment results were gathered in a "master" Microsoft Excel spreadsheet file (Figure 2) for the facilitation of storing information, the comparison of methodologies, and the derivation of meaningful conclusions for all the assessed criteria. In the master Excel file, the lines correspond to tools or to publications, while the columns correspond to the criteria mentioned above along with a summarised description of the tools and methods. Further information about the data used in each study and their availability were collected in a different Excel sheet, in which we focused mainly on the types of properties used as descriptors and on the availability of data. When the process of methodologies' analysis was completed, a different sheet focused on the quality assessment of the studies was created, where a quality score was assigned to each study based on the extracted information. The criteria used to perform the quality assessment of the methodologies were based on the key aspects of the analysis of the publications/tools (see §5.1.6).

Resource ID	Authors	Tool/method name	Description	Type of method	Endpoint	Endpoint type	Dataset size (samples)																																																																																																																																					
45	Papa et al.	SAR and QSAR models for cytotoxicity of TiO2 and ZnO	Using two empirical descriptors to develop local regression and classification models for TiO2 and ZnO NPs tested for their ability to disrupt the lipid membrane in cells. For classification authors used one cut-off value (i.e. LDH > 1.09) to distinguish between toxic (i.e. causing leaking or disruption of the membrane) and non-toxic effects due to exposure to TiO2 and ZnO NPs	Trained model	Membrane damage in cells measured by the release of lactate dehydrogenase (LDH)	numerical, categorical	42																																																																																																																																					
46	Walkey et al.	Cellular association prediction model of gold and silver NMs through PCF	Modeling of the biological interaction of the serum protein fingerprint related to cell association, via regression	Trained model	Cell association in A549 cell line (log2-transformed)	numerical	105																																																																																																																																					
47	Gajewicz et al.	Classification via the DF4nanoGrouping scheme	Using decision trees after reducing the number of descriptors from 285 to maximum three, to predict the surface activity of NMs in support of the DF4nanoGrouping framework (grouping system)	Trained model	Ferric reducing ability of serum (iOP), protein carbonylation, NOAEC from a short-term inhalation study in rats. Each endpoint is transformed in a binary class (active/inactive) according to reported thresholds	categorical	19																																																																																																																																					
48	Gajewicz et al.	Cytotoxicity prediction model of metal oxide nanoparticles [V]	Modeling of the toxicity as logarithmic values of molar 1/LC50 using two descriptors (enthalpy of formation and Mulliken's electronegativity)	Trained model	Cytotoxicity to human keratinocyte cell line (HaCat) expressed as the molar 1/log LC50	numerical	18																																																																																																																																					
49	Burello and Worth	Theoretical framework that predicts the oxidative stress potential of oxide nanoparticles	Theoretical model that predicts the oxidative stress potential of oxide NMs by comparing the redox potentials of relevant intracellular reactions with the energy structure, using reactivity descriptors.	Theoretical model	Oxidative stress potential	numerical	70																																																																																																																																					
50	Sayes and Ivanov	Cytotoxicity prediction of TiO2 and ZnO NMs	TiO2-Regression model to predict cellular membrane damage of TiO2 NMs via lactate dehydrogenase (LDH) release. ZnO-classification model to detect dense cell membranes and disrupted cell membrane after the exposure to ZnO NMs via LDH release.	Trained model	Cellular membrane damage via LDH release	numerical, categorical	24, 18																																																																																																																																					
51	Toropov et al.	Predictive modeling based on CORAL descriptors [J]	Application of the CORAL software to evaluate the Cytotoxicity to bacteria E. coli of a set of Metal Oxide NMs, using SMILES descriptors.	Trained model	Cytotoxicity to E. coli (logarithmic values of molar 1/EC50)	numerical	17																																																																																																																																					
52	<table><tr><th>Data used as independent variables (predictors)</th><th>Variable selection</th><th>Modelling approach</th><th>Type</th><th>Automation</th><th>Model validation</th><th>Statistics and metrics</th><th>Applicability domain definition</th><th>Variability assessment</th></tr><tr><td>Physicochemical descriptors, Conditions or medium related properties</td><td>Yes</td><td>MLR based on ordinary least squares (MLR-OLS), Support vector machine (SVM), Neural networks (RBFNN and GRNN), J48 classification tree</td><td>nanoQSAR</td><td>No</td><td>Internal and External validation</td><td>Pearson's correlation coefficient (r2), leave-one-out (Qloo2) and leave-many-out (Qlmo2) correlation coefficient, root mean square error (RMSE), accuracy, sensitivity, specificity</td><td>Yes</td><td>None reported</td></tr><tr><td>Physicochemical descriptors, In vitro data</td><td>Yes</td><td>Partial least-squares regression (PLSR)</td><td>nanoQSAR</td><td>No</td><td>Internal and external validation</td><td>Coefficient of determination (R2), CV-ANOVA (p-value), leave-one-out (Qloo2) and leave-many-out (Qlmo2) correlation coefficient</td><td>Yes</td><td>None reported</td></tr><tr><td>Physicochemical descriptors, Molecular properties</td><td>Yes</td><td>Decision tree models</td><td>nanoQSAR</td><td>No</td><td>External validation</td><td>Sensitivity, specificity, balanced accuracy, balanced error, confusion matrix</td><td>Yes</td><td>Multiple splitting of the dataset to training/test subsets</td></tr><tr><td>Molecular properties, Image descriptors</td><td>Yes</td><td>Multiple regression combined with a genetic algorithm (GA-MLR)</td><td>nano(QSAR)</td><td>No</td><td>Internal and external validation</td><td>leave-one-out cross-validation (CV-LOD), determination coefficient (R2), p-value, root mean square of cross-validation (RMSE)</td><td>Yes</td><td>Y-randomization</td></tr><tr><td>Molecular properties, Physicochemical descriptors</td><td>No</td><td>NA</td><td>Theoretical</td><td>No</td><td>None reported</td><td>R2</td><td>None reported</td><td>None reported</td></tr><tr><td>Physicochemical descriptors</td><td>Yes</td><td>Multivariate linear regression and linear discriminant analysis (LDA) based classification</td><td>nano(QSAR)</td><td>No</td><td>Internal validation</td><td>coefficient of determination, classification error estimated via resubstitution</td><td>None reported</td><td>None reported</td></tr><tr><td>Molecular properties</td><td>Yes</td><td colspan="7"><table><tr><th>Model interpretation</th><th>Dissemination</th><th>Model documentation and dissemination</th><th>Data gaps or missing steps</th><th>References</th><th>Year</th></tr><tr><td>Direct: mechanistic interpretation of the model, discussion about the limited number of the experimental data available and about the size descriptors' relevance</td><td>Scientific publication</td><td>No</td><td>Variability assessment is provided but not detailed. It is not stated if data are normalized prior to modelling.</td><td><a href="https://doi.org/10.1080/1062939X.2015.1080186">https://doi.org/10.1080/1062939X.2015.1080186</a></td><td>2015</td></tr><tr><td>Discussion on the proteins contributing to cell association</td><td>Scientific publication</td><td>No</td><td>Missing variability assessment. It is not stated if physicochemical descriptors are normalized</td><td><a href="https://doi.org/10.1021/nn406018q">https://doi.org/10.1021/nn406018q</a></td><td>2014</td></tr><tr><td>Direct: a mechanistic interpretation of the model is provided</td><td>Scientific publication</td><td>No</td><td>Insufficient metrics and statistics (accuracy or AUC/ROC missing), Applicability Domain via standardization of training set.</td><td><a href="https://doi.org/10.1080/17435590.2017.1415588">https://doi.org/10.1080/17435590.2017.1415588</a></td><td>2017</td></tr><tr><td>Direct: a mechanistic interpretation of the model is provided</td><td>Scientific publication</td><td>No</td><td></td><td><a href="https://doi.org/10.3109/17435590.2014.930195">https://doi.org/10.3109/17435590.2014.930195</a></td><td>2014</td></tr><tr><td>Direct: discussion on the toxicity mechanism of nanomaterials</td><td>Scientific publication</td><td>No</td><td>Missing statistics, it is not clear if R2 is correlation coefficient or correlation of determination</td><td><a href="https://doi.org/10.3109/17435590.2010.502980">https://doi.org/10.3109/17435590.2010.502980</a></td><td>2011</td></tr><tr><td>Direct: discussion on the descriptors</td><td>Scientific publication</td><td>No</td><td>Missing domain of applicability definition. More statistics needed</td><td><a href="https://doi.org/10.1111/j.1539-6924.2010.01488.x">https://doi.org/10.1111/j.1539-6924.2010.01488.x</a></td><td>2010</td></tr><tr><td>None reported</td><td>Scientific publication</td><td>No</td><td>Missing domain of applicability definition. It is not clear if R2 is squared correlation coefficient or coefficient of determination. Modelling methodology is unclear. Data normalization not performed</td><td><a href="http://dx.doi.org/10.1016/j.chemosphere.2012.05.077">http://dx.doi.org/10.1016/j.chemosphere.2012.05.077</a></td><td>2012</td></tr><tr><td>None reported</td><td>Scientific publication, development of a web application: <a href="http://padel.nus.edu.sg/software/padelidspredictor/">http://padel.nus.edu.sg/software/padelidspredictor/</a></td><td>No</td><td>It is not stated if physicochemical descriptors are normalized. Applicability domain via Multiple threshold method</td><td><a href="https://doi.org/10.1039/C2RA21489J">https://doi.org/10.1039/C2RA21489J</a></td><td>2012</td></tr><tr><td>Direct: the QSTR-perturbation model gives deeper insights of the overall toxicological behavior of NPs, playing also an important role in their virtual screening for both ecotoxicity and cytotoxicity</td><td>Scientific publication</td><td>No</td><td>Missing domain applicability definition, data not normalized, train-test not provided</td><td><a href="https://doi.org/10.1021/es503861x">https://doi.org/10.1021/es503861x</a></td><td>2014</td></tr></table></td></tr></table>								Data used as independent variables (predictors)	Variable selection	Modelling approach	Type	Automation	Model validation	Statistics and metrics	Applicability domain definition	Variability assessment	Physicochemical descriptors, Conditions or medium related properties	Yes	MLR based on ordinary least squares (MLR-OLS), Support vector machine (SVM), Neural networks (RBFNN and GRNN), J48 classification tree	nanoQSAR	No	Internal and External validation	Pearson's correlation coefficient (r2), leave-one-out (Qloo2) and leave-many-out (Qlmo2) correlation coefficient, root mean square error (RMSE), accuracy, sensitivity, specificity	Yes	None reported	Physicochemical descriptors, In vitro data	Yes	Partial least-squares regression (PLSR)	nanoQSAR	No	Internal and external validation	Coefficient of determination (R2), CV-ANOVA (p-value), leave-one-out (Qloo2) and leave-many-out (Qlmo2) correlation coefficient	Yes	None reported	Physicochemical descriptors, Molecular properties	Yes	Decision tree models	nanoQSAR	No	External validation	Sensitivity, specificity, balanced accuracy, balanced error, confusion matrix	Yes	Multiple splitting of the dataset to training/test subsets	Molecular properties, Image descriptors	Yes	Multiple regression combined with a genetic algorithm (GA-MLR)	nano(QSAR)	No	Internal and external validation	leave-one-out cross-validation (CV-LOD), determination coefficient (R2), p-value, root mean square of cross-validation (RMSE)	Yes	Y-randomization	Molecular properties, Physicochemical descriptors	No	NA	Theoretical	No	None reported	R2	None reported	None reported	Physicochemical descriptors	Yes	Multivariate linear regression and linear discriminant analysis (LDA) based classification	nano(QSAR)	No	Internal validation	coefficient of determination, classification error estimated via resubstitution	None reported	None reported	Molecular properties	Yes	<table><tr><th>Model interpretation</th><th>Dissemination</th><th>Model documentation and dissemination</th><th>Data gaps or missing steps</th><th>References</th><th>Year</th></tr><tr><td>Direct: mechanistic interpretation of the model, discussion about the limited number of the experimental data available and about the size descriptors' relevance</td><td>Scientific publication</td><td>No</td><td>Variability assessment is provided but not detailed. It is not stated if data are normalized prior to modelling.</td><td><a href="https://doi.org/10.1080/1062939X.2015.1080186">https://doi.org/10.1080/1062939X.2015.1080186</a></td><td>2015</td></tr><tr><td>Discussion on the proteins contributing to cell association</td><td>Scientific publication</td><td>No</td><td>Missing variability assessment. It is not stated if physicochemical descriptors are normalized</td><td><a href="https://doi.org/10.1021/nn406018q">https://doi.org/10.1021/nn406018q</a></td><td>2014</td></tr><tr><td>Direct: a mechanistic interpretation of the model is provided</td><td>Scientific publication</td><td>No</td><td>Insufficient metrics and statistics (accuracy or AUC/ROC missing), Applicability Domain via standardization of training set.</td><td><a href="https://doi.org/10.1080/17435590.2017.1415588">https://doi.org/10.1080/17435590.2017.1415588</a></td><td>2017</td></tr><tr><td>Direct: a mechanistic interpretation of the model is provided</td><td>Scientific publication</td><td>No</td><td></td><td><a href="https://doi.org/10.3109/17435590.2014.930195">https://doi.org/10.3109/17435590.2014.930195</a></td><td>2014</td></tr><tr><td>Direct: discussion on the toxicity mechanism of nanomaterials</td><td>Scientific publication</td><td>No</td><td>Missing statistics, it is not clear if R2 is correlation coefficient or correlation of determination</td><td><a href="https://doi.org/10.3109/17435590.2010.502980">https://doi.org/10.3109/17435590.2010.502980</a></td><td>2011</td></tr><tr><td>Direct: discussion on the descriptors</td><td>Scientific publication</td><td>No</td><td>Missing domain of applicability definition. More statistics needed</td><td><a href="https://doi.org/10.1111/j.1539-6924.2010.01488.x">https://doi.org/10.1111/j.1539-6924.2010.01488.x</a></td><td>2010</td></tr><tr><td>None reported</td><td>Scientific publication</td><td>No</td><td>Missing domain of applicability definition. It is not clear if R2 is squared correlation coefficient or coefficient of determination. Modelling methodology is unclear. Data normalization not performed</td><td><a href="http://dx.doi.org/10.1016/j.chemosphere.2012.05.077">http://dx.doi.org/10.1016/j.chemosphere.2012.05.077</a></td><td>2012</td></tr><tr><td>None reported</td><td>Scientific publication, development of a web application: <a href="http://padel.nus.edu.sg/software/padelidspredictor/">http://padel.nus.edu.sg/software/padelidspredictor/</a></td><td>No</td><td>It is not stated if physicochemical descriptors are normalized. Applicability domain via Multiple threshold method</td><td><a href="https://doi.org/10.1039/C2RA21489J">https://doi.org/10.1039/C2RA21489J</a></td><td>2012</td></tr><tr><td>Direct: the QSTR-perturbation model gives deeper insights of the overall toxicological behavior of NPs, playing also an important role in their virtual screening for both ecotoxicity and cytotoxicity</td><td>Scientific publication</td><td>No</td><td>Missing domain applicability definition, data not normalized, train-test not provided</td><td><a href="https://doi.org/10.1021/es503861x">https://doi.org/10.1021/es503861x</a></td><td>2014</td></tr></table>							Model interpretation	Dissemination	Model documentation and dissemination	Data gaps or missing steps	References	Year	Direct: mechanistic interpretation of the model, discussion about the limited number of the experimental data available and about the size descriptors' relevance	Scientific publication	No	Variability assessment is provided but not detailed. It is not stated if data are normalized prior to modelling.	<a href="https://doi.org/10.1080/1062939X.2015.1080186">https://doi.org/10.1080/1062939X.2015.1080186</a>	2015	Discussion on the proteins contributing to cell association	Scientific publication	No	Missing variability assessment. It is not stated if physicochemical descriptors are normalized	<a href="https://doi.org/10.1021/nn406018q">https://doi.org/10.1021/nn406018q</a>	2014	Direct: a mechanistic interpretation of the model is provided	Scientific publication	No	Insufficient metrics and statistics (accuracy or AUC/ROC missing), Applicability Domain via standardization of training set.	<a href="https://doi.org/10.1080/17435590.2017.1415588">https://doi.org/10.1080/17435590.2017.1415588</a>	2017	Direct: a mechanistic interpretation of the model is provided	Scientific publication	No		<a href="https://doi.org/10.3109/17435590.2014.930195">https://doi.org/10.3109/17435590.2014.930195</a>	2014	Direct: discussion on the toxicity mechanism of nanomaterials	Scientific publication	No	Missing statistics, it is not clear if R2 is correlation coefficient or correlation of determination	<a href="https://doi.org/10.3109/17435590.2010.502980">https://doi.org/10.3109/17435590.2010.502980</a>	2011	Direct: discussion on the descriptors	Scientific publication	No	Missing domain of applicability definition. More statistics needed	<a href="https://doi.org/10.1111/j.1539-6924.2010.01488.x">https://doi.org/10.1111/j.1539-6924.2010.01488.x</a>	2010	None reported	Scientific publication	No	Missing domain of applicability definition. It is not clear if R2 is squared correlation coefficient or coefficient of determination. Modelling methodology is unclear. Data normalization not performed	<a href="http://dx.doi.org/10.1016/j.chemosphere.2012.05.077">http://dx.doi.org/10.1016/j.chemosphere.2012.05.077</a>	2012	None reported	Scientific publication, development of a web application: <a href="http://padel.nus.edu.sg/software/padelidspredictor/">http://padel.nus.edu.sg/software/padelidspredictor/</a>	No	It is not stated if physicochemical descriptors are normalized. Applicability domain via Multiple threshold method	<a href="https://doi.org/10.1039/C2RA21489J">https://doi.org/10.1039/C2RA21489J</a>	2012	Direct: the QSTR-perturbation model gives deeper insights of the overall toxicological behavior of NPs, playing also an important role in their virtual screening for both ecotoxicity and cytotoxicity	Scientific publication	No	Missing domain applicability definition, data not normalized, train-test not provided	<a href="https://doi.org/10.1021/es503861x">https://doi.org/10.1021/es503861x</a>	2014
Data used as independent variables (predictors)	Variable selection	Modelling approach	Type	Automation	Model validation	Statistics and metrics	Applicability domain definition	Variability assessment																																																																																																																																				
Physicochemical descriptors, Conditions or medium related properties	Yes	MLR based on ordinary least squares (MLR-OLS), Support vector machine (SVM), Neural networks (RBFNN and GRNN), J48 classification tree	nanoQSAR	No	Internal and External validation	Pearson's correlation coefficient (r2), leave-one-out (Qloo2) and leave-many-out (Qlmo2) correlation coefficient, root mean square error (RMSE), accuracy, sensitivity, specificity	Yes	None reported																																																																																																																																				
Physicochemical descriptors, In vitro data	Yes	Partial least-squares regression (PLSR)	nanoQSAR	No	Internal and external validation	Coefficient of determination (R2), CV-ANOVA (p-value), leave-one-out (Qloo2) and leave-many-out (Qlmo2) correlation coefficient	Yes	None reported																																																																																																																																				
Physicochemical descriptors, Molecular properties	Yes	Decision tree models	nanoQSAR	No	External validation	Sensitivity, specificity, balanced accuracy, balanced error, confusion matrix	Yes	Multiple splitting of the dataset to training/test subsets																																																																																																																																				
Molecular properties, Image descriptors	Yes	Multiple regression combined with a genetic algorithm (GA-MLR)	nano(QSAR)	No	Internal and external validation	leave-one-out cross-validation (CV-LOD), determination coefficient (R2), p-value, root mean square of cross-validation (RMSE)	Yes	Y-randomization																																																																																																																																				
Molecular properties, Physicochemical descriptors	No	NA	Theoretical	No	None reported	R2	None reported	None reported																																																																																																																																				
Physicochemical descriptors	Yes	Multivariate linear regression and linear discriminant analysis (LDA) based classification	nano(QSAR)	No	Internal validation	coefficient of determination, classification error estimated via resubstitution	None reported	None reported																																																																																																																																				
Molecular properties	Yes	<table><tr><th>Model interpretation</th><th>Dissemination</th><th>Model documentation and dissemination</th><th>Data gaps or missing steps</th><th>References</th><th>Year</th></tr><tr><td>Direct: mechanistic interpretation of the model, discussion about the limited number of the experimental data available and about the size descriptors' relevance</td><td>Scientific publication</td><td>No</td><td>Variability assessment is provided but not detailed. It is not stated if data are normalized prior to modelling.</td><td><a href="https://doi.org/10.1080/1062939X.2015.1080186">https://doi.org/10.1080/1062939X.2015.1080186</a></td><td>2015</td></tr><tr><td>Discussion on the proteins contributing to cell association</td><td>Scientific publication</td><td>No</td><td>Missing variability assessment. It is not stated if physicochemical descriptors are normalized</td><td><a href="https://doi.org/10.1021/nn406018q">https://doi.org/10.1021/nn406018q</a></td><td>2014</td></tr><tr><td>Direct: a mechanistic interpretation of the model is provided</td><td>Scientific publication</td><td>No</td><td>Insufficient metrics and statistics (accuracy or AUC/ROC missing), Applicability Domain via standardization of training set.</td><td><a href="https://doi.org/10.1080/17435590.2017.1415588">https://doi.org/10.1080/17435590.2017.1415588</a></td><td>2017</td></tr><tr><td>Direct: a mechanistic interpretation of the model is provided</td><td>Scientific publication</td><td>No</td><td></td><td><a href="https://doi.org/10.3109/17435590.2014.930195">https://doi.org/10.3109/17435590.2014.930195</a></td><td>2014</td></tr><tr><td>Direct: discussion on the toxicity mechanism of nanomaterials</td><td>Scientific publication</td><td>No</td><td>Missing statistics, it is not clear if R2 is correlation coefficient or correlation of determination</td><td><a href="https://doi.org/10.3109/17435590.2010.502980">https://doi.org/10.3109/17435590.2010.502980</a></td><td>2011</td></tr><tr><td>Direct: discussion on the descriptors</td><td>Scientific publication</td><td>No</td><td>Missing domain of applicability definition. More statistics needed</td><td><a href="https://doi.org/10.1111/j.1539-6924.2010.01488.x">https://doi.org/10.1111/j.1539-6924.2010.01488.x</a></td><td>2010</td></tr><tr><td>None reported</td><td>Scientific publication</td><td>No</td><td>Missing domain of applicability definition. It is not clear if R2 is squared correlation coefficient or coefficient of determination. Modelling methodology is unclear. Data normalization not performed</td><td><a href="http://dx.doi.org/10.1016/j.chemosphere.2012.05.077">http://dx.doi.org/10.1016/j.chemosphere.2012.05.077</a></td><td>2012</td></tr><tr><td>None reported</td><td>Scientific publication, development of a web application: <a href="http://padel.nus.edu.sg/software/padelidspredictor/">http://padel.nus.edu.sg/software/padelidspredictor/</a></td><td>No</td><td>It is not stated if physicochemical descriptors are normalized. Applicability domain via Multiple threshold method</td><td><a href="https://doi.org/10.1039/C2RA21489J">https://doi.org/10.1039/C2RA21489J</a></td><td>2012</td></tr><tr><td>Direct: the QSTR-perturbation model gives deeper insights of the overall toxicological behavior of NPs, playing also an important role in their virtual screening for both ecotoxicity and cytotoxicity</td><td>Scientific publication</td><td>No</td><td>Missing domain applicability definition, data not normalized, train-test not provided</td><td><a href="https://doi.org/10.1021/es503861x">https://doi.org/10.1021/es503861x</a></td><td>2014</td></tr></table>							Model interpretation	Dissemination	Model documentation and dissemination	Data gaps or missing steps	References	Year	Direct: mechanistic interpretation of the model, discussion about the limited number of the experimental data available and about the size descriptors' relevance	Scientific publication	No	Variability assessment is provided but not detailed. It is not stated if data are normalized prior to modelling.	<a href="https://doi.org/10.1080/1062939X.2015.1080186">https://doi.org/10.1080/1062939X.2015.1080186</a>	2015	Discussion on the proteins contributing to cell association	Scientific publication	No	Missing variability assessment. It is not stated if physicochemical descriptors are normalized	<a href="https://doi.org/10.1021/nn406018q">https://doi.org/10.1021/nn406018q</a>	2014	Direct: a mechanistic interpretation of the model is provided	Scientific publication	No	Insufficient metrics and statistics (accuracy or AUC/ROC missing), Applicability Domain via standardization of training set.	<a href="https://doi.org/10.1080/17435590.2017.1415588">https://doi.org/10.1080/17435590.2017.1415588</a>	2017	Direct: a mechanistic interpretation of the model is provided	Scientific publication	No		<a href="https://doi.org/10.3109/17435590.2014.930195">https://doi.org/10.3109/17435590.2014.930195</a>	2014	Direct: discussion on the toxicity mechanism of nanomaterials	Scientific publication	No	Missing statistics, it is not clear if R2 is correlation coefficient or correlation of determination	<a href="https://doi.org/10.3109/17435590.2010.502980">https://doi.org/10.3109/17435590.2010.502980</a>	2011	Direct: discussion on the descriptors	Scientific publication	No	Missing domain of applicability definition. More statistics needed	<a href="https://doi.org/10.1111/j.1539-6924.2010.01488.x">https://doi.org/10.1111/j.1539-6924.2010.01488.x</a>	2010	None reported	Scientific publication	No	Missing domain of applicability definition. It is not clear if R2 is squared correlation coefficient or coefficient of determination. Modelling methodology is unclear. Data normalization not performed	<a href="http://dx.doi.org/10.1016/j.chemosphere.2012.05.077">http://dx.doi.org/10.1016/j.chemosphere.2012.05.077</a>	2012	None reported	Scientific publication, development of a web application: <a href="http://padel.nus.edu.sg/software/padelidspredictor/">http://padel.nus.edu.sg/software/padelidspredictor/</a>	No	It is not stated if physicochemical descriptors are normalized. Applicability domain via Multiple threshold method	<a href="https://doi.org/10.1039/C2RA21489J">https://doi.org/10.1039/C2RA21489J</a>	2012	Direct: the QSTR-perturbation model gives deeper insights of the overall toxicological behavior of NPs, playing also an important role in their virtual screening for both ecotoxicity and cytotoxicity	Scientific publication	No	Missing domain applicability definition, data not normalized, train-test not provided	<a href="https://doi.org/10.1021/es503861x">https://doi.org/10.1021/es503861x</a>	2014																																																																								
Model interpretation	Dissemination	Model documentation and dissemination	Data gaps or missing steps	References	Year																																																																																																																																							
Direct: mechanistic interpretation of the model, discussion about the limited number of the experimental data available and about the size descriptors' relevance	Scientific publication	No	Variability assessment is provided but not detailed. It is not stated if data are normalized prior to modelling.	<a href="https://doi.org/10.1080/1062939X.2015.1080186">https://doi.org/10.1080/1062939X.2015.1080186</a>	2015																																																																																																																																							
Discussion on the proteins contributing to cell association	Scientific publication	No	Missing variability assessment. It is not stated if physicochemical descriptors are normalized	<a href="https://doi.org/10.1021/nn406018q">https://doi.org/10.1021/nn406018q</a>	2014																																																																																																																																							
Direct: a mechanistic interpretation of the model is provided	Scientific publication	No	Insufficient metrics and statistics (accuracy or AUC/ROC missing), Applicability Domain via standardization of training set.	<a href="https://doi.org/10.1080/17435590.2017.1415588">https://doi.org/10.1080/17435590.2017.1415588</a>	2017																																																																																																																																							
Direct: a mechanistic interpretation of the model is provided	Scientific publication	No		<a href="https://doi.org/10.3109/17435590.2014.930195">https://doi.org/10.3109/17435590.2014.930195</a>	2014																																																																																																																																							
Direct: discussion on the toxicity mechanism of nanomaterials	Scientific publication	No	Missing statistics, it is not clear if R2 is correlation coefficient or correlation of determination	<a href="https://doi.org/10.3109/17435590.2010.502980">https://doi.org/10.3109/17435590.2010.502980</a>	2011																																																																																																																																							
Direct: discussion on the descriptors	Scientific publication	No	Missing domain of applicability definition. More statistics needed	<a href="https://doi.org/10.1111/j.1539-6924.2010.01488.x">https://doi.org/10.1111/j.1539-6924.2010.01488.x</a>	2010																																																																																																																																							
None reported	Scientific publication	No	Missing domain of applicability definition. It is not clear if R2 is squared correlation coefficient or coefficient of determination. Modelling methodology is unclear. Data normalization not performed	<a href="http://dx.doi.org/10.1016/j.chemosphere.2012.05.077">http://dx.doi.org/10.1016/j.chemosphere.2012.05.077</a>	2012																																																																																																																																							
None reported	Scientific publication, development of a web application: <a href="http://padel.nus.edu.sg/software/padelidspredictor/">http://padel.nus.edu.sg/software/padelidspredictor/</a>	No	It is not stated if physicochemical descriptors are normalized. Applicability domain via Multiple threshold method	<a href="https://doi.org/10.1039/C2RA21489J">https://doi.org/10.1039/C2RA21489J</a>	2012																																																																																																																																							
Direct: the QSTR-perturbation model gives deeper insights of the overall toxicological behavior of NPs, playing also an important role in their virtual screening for both ecotoxicity and cytotoxicity	Scientific publication	No	Missing domain applicability definition, data not normalized, train-test not provided	<a href="https://doi.org/10.1021/es503861x">https://doi.org/10.1021/es503861x</a>	2014																																																																																																																																							

Figure 2: Systematic collection of information from the reviewed publications and tools in the Excel file.

## 5 Critical analysis of the collected *in silico* tools and available models

In this part of the report, we are presenting the results of the reviewed publications and models in respect to the methodologies of nanoQSAR, read-across, PBPK, AOP and the simulations (Table 2). The main part of the review is dedicated to the nanoQSAR and grouping/read-across models as they are considered the primarily data-driven *in silico* methods for the risk assessment of NMs. The main key points of the available models are included.

Table 2: Number of assessed studies for different modelling methodologies.

Type of computational methodology	Number of studies
nanoQSARs and grouping/read-across	132
PBPKs	11
MD simulations	28
AOPs	8
Other models (exposure, causal, dose-response)	11

The selection of an *in silico* methodology over another to support the NMs hazard and risk assessment is a question of data and expedient outcomes. In Table 3 the required input and expected output information is presented for each computational methodology, as well as their advantages and disadvantages. The methodologies' limitations related to computing power might be eliminated in the future as more potent hardware (e.g., graphics processing unit, GPU, accelerated calculations) and optimised software (e.g., parallelisation of processes) are being developed. It is also noted that in any selected computational methodology for the study of the NMs properties or interactions with biological systems, experimental confirmation of the results may still be needed.



Table 3: Comparison of the assessed *in silico* methodologies included in this project.

Methodology	Type of data	Application	Advantages	Disadvantages
<b>ML based methods</b>	Typically -for a collection of NM samples- a set of properties/descriptors (used as independent variables) and a target/endpoint variable (dependent variable), in tabular format.	<ul style="list-style-type: none"> <li>Supervised learning: Predict a variable of interest (e.g., NMs toxicity-related endpoints, physicochemical properties, etc.).</li> <li>Unsupervised learning: Detect patterns in unlabelled datasets, clustering (e.g., for grouping of similar NMs).</li> </ul>	<ul style="list-style-type: none"> <li>They can handle multi-dimensional data and they can be used to model different endpoints (e.g., NMs toxicity or other properties).</li> <li>The ML and the statistical algorithms can reveal interesting trends and highlight important descriptors that influence the endpoint of interest, only based on the information that hold input data.</li> <li>Some ML algorithms support missing values (e.g., appropriate when NM data are scarce).</li> <li>They can be easily automated to minimise stakeholders (especially non-informatics experts) interaction with the models.</li> <li>They can be easily used for NMs screening to prioritise favourable candidates for synthesis and experimental evaluation.</li> </ul>	<ul style="list-style-type: none"> <li>Models trained on experimental or theoretical data of low quality propagate errors, produce false correlations, and may lead to interpretation inaccuracies (even if the model is considered to produce reliable predictions).</li> <li>The models may lack “physical reality” if emphasis is given solely on the development of accurate, robust models based on statistical inference and sophisticated ML algorithms. The goal of the <i>in silico</i>-supported NMs risk assessment is to use the models as tools to interpret underlying physical or biological phenomena (e.g., NMs toxicity), and not to develop mere data-driven models.</li> <li>Significantly large datasets are required to develop accurate models.</li> <li>They risk being outdated if they are not frequently updated with newly produced data.</li> </ul>
nanoQSAR			<ul style="list-style-type: none"> <li>They quantify the relationship of NMs structural properties to the query endpoint.</li> <li>They are expanded to consider NM specificities (physicochemical properties, theoretical, molecular, and structural descriptors of the NMs core and coating).</li> </ul>	<ul style="list-style-type: none"> <li>Large datasets must be provided so that QSAR-type models are functional (limited nano-related datasets).</li> <li>A common mechanism of toxicity is a prerequisite for modelling (the NMs variety and all the possible transformations that may occur during the NMs life cycle induce different toxicity mechanisms)<sup>6,26</sup>.</li> </ul>

Grouping and read-across			<ul style="list-style-type: none"> <li>• Can be applied in cases of limited data (e.g., NM datasets with less than 20 samples).</li> <li>• They can handle data of different types (e.g., physicochemical, image, periodic-table, theoretical descriptors, toxicogenomics data, etc.).</li> </ul>	<ul style="list-style-type: none"> <li>• Difficulties in the perception of similarity between NMs.</li> <li>• The lack of knowledge of the presumed mode-of-action driving the endpoint of interest.</li> <li>• Read-across does not necessarily imply the development of a model (e.g., development of conceptual frameworks without any practical information on assays and thresholds)<sup>27</sup>.</li> </ul>
Deep learning	Big data (e.g., microarray data), unstructured data (e.g., NMs TEM images).		<ul style="list-style-type: none"> <li>• Features are learned directly from the data.</li> </ul>	<ul style="list-style-type: none"> <li>• A very large amount of data is required to build the model.</li> <li>• It is a computationally expensive methodology.</li> <li>• In case of too complex models (many hidden layers) they are difficult to interpret.</li> </ul>
PBPK modelling	System's (organism) data (e.g., weight, age, genetics, etc.), NM related data (e.g., physicochemical properties), and external exposure data (e.g., administration route, dose, or exposure duration).	Prediction of internal distribution of NMs in an organism after exposure.	<ul style="list-style-type: none"> <li>• Ability to perform inter-species (e.g., from rat to human) or intra-species (e.g., from adults to children) extrapolation through scaling methods.</li> <li>• Integration of literature physiology data, contributes to the limitation of clinical trials.</li> </ul>	<ul style="list-style-type: none"> <li>• Validation of PBPK models requires extensive <i>in vitro</i> and <i>in vivo</i> experimental data<sup>28</sup>.</li> <li>• Validation of PBPK models may be challenging in cases of inter-species extrapolation where relevant (human) biodistribution data are not available<sup>29</sup>.</li> <li>• The development of PBPK models based on <i>in vitro</i> data should be performed with caution, as NMs when entering the organism acquire a new biological identity (e.g., biocorona formation, altered surface chemistry, etc.)<sup>8</sup> or undergo different transformations (e.g., dissolution)<sup>19</sup>.</li> </ul>

MD simulations	Size and composition of the NMs and composition of the NM coating (if existent). Information on the studied system (e.g., proteins, lipid membranes, etc.), pressure, temperature, or volume of the simulation box. Force field.	Understanding of the mechanism of the interaction of different types of NMs with biological systems using classical mechanics and prediction of macroscopic NM properties.	<ul style="list-style-type: none"> <li>• The gained knowledge on the nano-bio interactions using MD (e.g., underlying mechanism) cannot be easily obtained experimentally.</li> <li>• As virtual experiments are designed from scratch there is a possibility of simulating systems that are not extensively studied experimentally (e.g., graphene in biological environments).</li> <li>• The possible toxicological effects induced by the interaction between NMs and biomolecules can be predicted and integrated in the SbD of novel NMs.</li> </ul>	<ul style="list-style-type: none"> <li>• The high computational cost of these methodologies allows analysis only for small systems and for limited time scales.</li> <li>• To speed up simulations coarse-grained force fields can be used (instead of density functional theory simulations), that may affect the accuracy of the results.</li> </ul>
AOPs	Any biological data (including toxicogenomics). Any <i>in vivo</i> , <i>in vitro</i> and <i>in silico</i> data to support the AOP DoA.	Description of the molecular mechanisms that lead to adverse outcomes (e.g., lung fibrosis) starting from a molecular initiating event (e.g., occupational exposure to NMs inducing interaction with the lung resident cell membrane components, AOP: 173) <sup>24</sup> .	<ul style="list-style-type: none"> <li>• Systematic collection of weight of evidence and knowledge organisation.</li> <li>• Link of mechanistic endpoints (e.g., DNA mutation) to apical endpoints (e.g., carcinogenesis). Use of alternative-fast tests in NMs risk assessment.</li> <li>• Read-across similarities between NMs with respect to AOP activation can be determined (in case of available quantitative data)<sup>30</sup>.</li> <li>• Identification of knowledge gaps<sup>8</sup>.</li> </ul> <p>More information can be found in Leist <i>et al.</i><sup>30</sup>.</p>	<ul style="list-style-type: none"> <li>• The effect of time cannot be considered in the AOPs.</li> <li>• Binning of large groups of events<sup>30</sup>.</li> <li>• Several AO-related limitations (e.g., for inflammation-mediated AOPs: the complexity of measuring and quantifying inflammation <i>in vitro</i>)<sup>31</sup>.</li> </ul> <p>More information can be found in Leist <i>et al.</i><sup>30</sup>.</p>

## 5.1 nanoQSAR/read-across

In Appendix I, a list of all the assessed grouping/read-across and nanoQSAR studies and models is presented, including all the key information of each model (modelling technique, validation, applicability domain, data availability, etc.).

### 5.1.1 Data and availability

In order to study and analyse the entire life cycle of NMs and to create *in silico* approaches as alternatives to costly and time-consuming experimental procedures for the assessment of NMs toxicity, the scientific community largely relies on data. Data are either produced by the scientific teams that later develop the predictive models, or they are retrieved from the literature or NMs-specific databases (Figure 3).

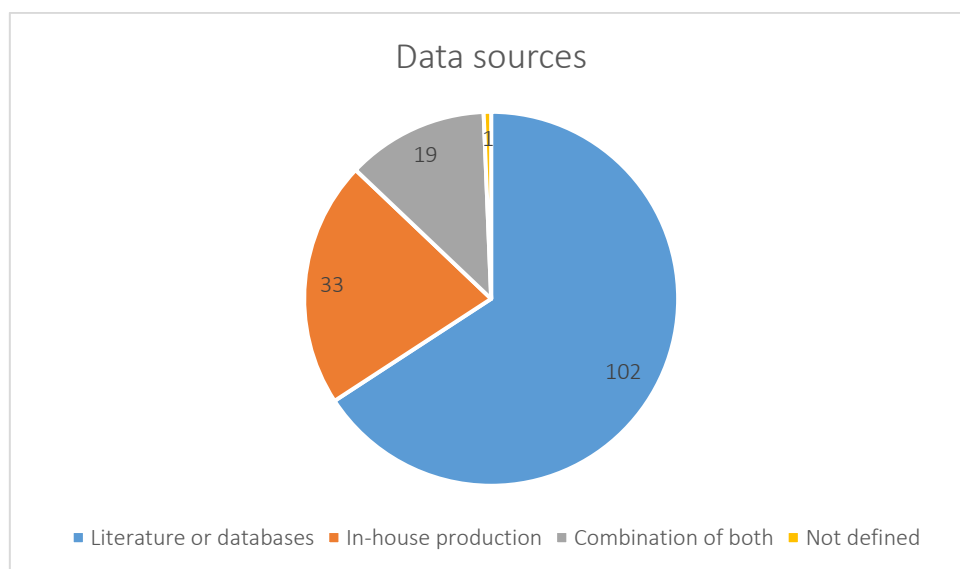


Figure 3: Data sources for the datasets used to build nanoQSAR and read-across models. Some studies included more than one model and thus more datasets were used in the same study.

Furthermore, data scarcity is a known issue when studying the NMs entire life cycle<sup>32</sup>. In this review we also recorded the studies that were based on publicly available datasets (e.g., datasets included in databases, in the supporting information of scientific publications, or they are produced by the same scientific team performing the modelling). Data unavailability was marked as a result, even in cases that only a part of the data is available (e.g., toxicity endpoint available but molecular descriptors used in modelling not available). The results are presented schematically in Figure 4.

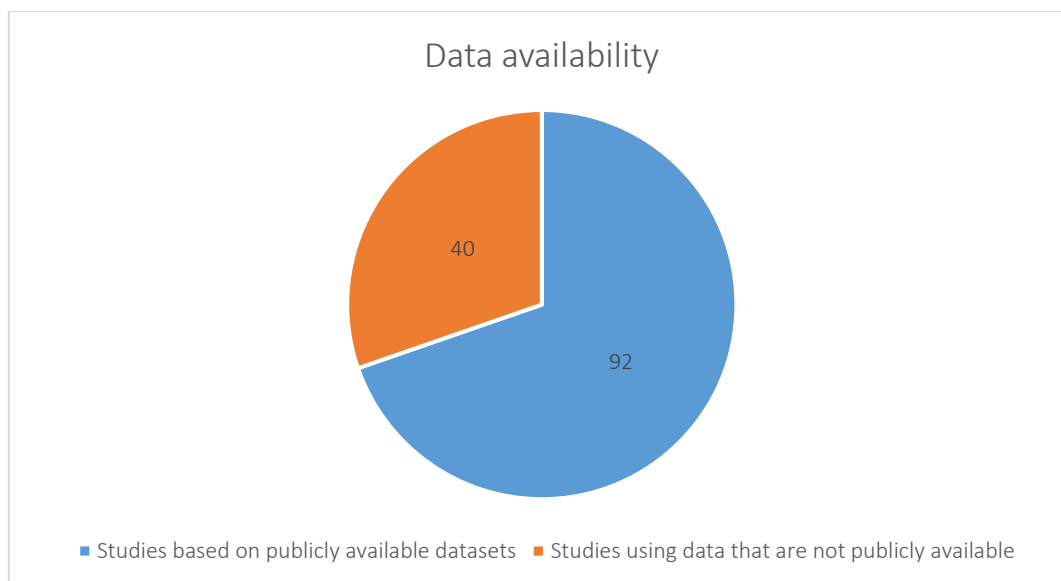


Figure 4: nanoQSAR and read-across models trained with publicly available data.

The datasets used regularly in the development of nanoQSAR and read-across predictive models and methodologies are presented below. The number of tools/methodologies that were based on these or part of these datasets is presented in the brackets.

- This dataset, generated by Walkey *et al.*<sup>33</sup>, consists of 105 gold (Au) and 16 silver (Ag) NPs with different surface functionalisation. The organic surface ligands are classified according to their chemical structure and net charge at physiological pH, as “neutral”, “anionic” and “cationic”. NPs physicochemical properties as well as their protein corona composition (fingerprint) are available. The endpoint of interest is the NPs cell association to A549 cell line, considering the magnesium content of the cells. The cell association values are log<sub>2</sub>-transformed prior to modelling. [#studies: 9<sup>33–41</sup>],
- Weissleder *et al.*<sup>42</sup> synthesised a library of 146 NPs with the same metal core (a monocrystalline magnetic NP, with a 3 nm core of (Fe<sub>2</sub>O<sub>3</sub>)<sub>n</sub>(Fe<sub>3</sub>O<sub>4</sub>)<sub>m</sub> covered with a layer of 10 kDa dextran, that was cross-linked with epichlorohydrin and aminated by reaction with NH<sub>3</sub>) and different surface modifiers (organic small molecules conjugated to the NMs surface). The cellular uptake of the NPs in human pancreatic cancer (PaCa2) cells (endpoint) was measured and the respective values are presented for 109 of these NPs in the publication of Fourches *et al.*<sup>43</sup>. Considering that the NPs have the same metal core, they are characterised by the surface-modifying molecules and thus, cheminformatics-based descriptors can be generated for each NP complex. Different research groups calculate molecular descriptors based on the surface organic molecules’ SMILES specification by employing different methodologies and use them as input data to the developed models. [#studies: 7<sup>43–49</sup>],
- Puzyn *et al.*<sup>50</sup> developed a dataset of 17 metal oxide (MeOx) NMs (ZnO, CuO, Al<sub>2</sub>O<sub>3</sub>, Fe<sub>2</sub>O<sub>3</sub>, SnO<sub>2</sub>, TiO<sub>2</sub>, V<sub>2</sub>O<sub>3</sub>, Y<sub>2</sub>O<sub>3</sub>, Bi<sub>2</sub>O<sub>3</sub>, In<sub>2</sub>O<sub>3</sub>, Sb<sub>2</sub>O<sub>3</sub>, SiO<sub>2</sub>, ZrO<sub>2</sub>, CoO, NiO, Cr<sub>2</sub>O<sub>3</sub>, and La<sub>2</sub>O<sub>3</sub>) assessed for their cytotoxicity to bacteria *Escherichia coli*. Cytotoxicity is expressed in logarithmic values of molar 1/EC<sub>50</sub> (the effective concentration of a given MeOx that reduces bacterial viability by 50%). Semi-empirical quantum-mechanical calculations produced molecular parameters that encode NMs reactivity-related electronic properties. [#studies: 15<sup>26,39,58–62,50–57</sup>],
- This dataset, generated by Gajewicz *et al.*<sup>63</sup>, consists of 18 MeOx NPs (ZnO, Al<sub>2</sub>O<sub>3</sub>, Fe<sub>2</sub>O<sub>3</sub>, SnO<sub>2</sub>, TiO<sub>2</sub>, V<sub>2</sub>O<sub>3</sub>, Y<sub>2</sub>O<sub>3</sub>, Bi<sub>2</sub>O<sub>3</sub>, In<sub>2</sub>O<sub>3</sub>, Sb<sub>2</sub>O<sub>3</sub>, SiO<sub>2</sub>, ZrO<sub>2</sub>, CoO, NiO, Cr<sub>2</sub>O<sub>3</sub>, WO<sub>3</sub>, Mn<sub>2</sub>O<sub>3</sub> and La<sub>2</sub>O<sub>3</sub>) tested for their toxicity to the human keratinocyte cell line (HaCaT). Cell viability was measured and expressed in logarithmic values of molar 1/LC<sub>50</sub> (the concentration that provoked a 50% reduction of the cells after 24 h exposure). For the NPs quantum-mechanical -calculated with

semi-empirical methods- and image descriptors -calculated from NPs TEM images- are provided and can be used as input variables to the *in silico* models. [#studies: 12<sup>26,39,62,63,49,51,54-57,59,61</sup>],

- Shaw *et al*<sup>64</sup>. developed a dataset originally consisting of 51 coated iron oxide NPs, determining their toxicity class as either “active” or “inactive”. Activity was described by a “bioactivity profile” which averaged 64 features and conditions based on the assay z-scores for all possible combinations of 4 doses, 4 different cell lines, and 4 different *in vitro* assays. [#studies: 3<sup>43,65,66</sup> where authors used 44 out of 51 NPs],
- Zhou *et al*<sup>67</sup>. generated a dataset consisting of 83 functionalised multi-walled carbon nanotubes (MWCNTs) experimentally characterised using six *in vitro* assays with surface molecular diversity design, combinatorial library synthesis, and multiple biological screenings. The 6 assays were: protein binding to proteins bovine serum albumin (BSA), carbonic anhydrase (CA), chymotrypsin (CT), and haemoglobin (HB), nitrogen oxide (NO) generation and cell viability by WST-1 assay. Again, the MWCNTs are characterised by their surface-modifying molecules and thus, cheminformatics-based descriptors can be generated for each MWCNTs complex based on the surface organic molecules’ SMILES specification. Different research groups calculate molecular descriptors by employing different methodologies and use them as input data to the developed models. [#studies: 3<sup>68-70</sup>].
- In the publication of Oksel *et al*.<sup>71</sup> additional datasets used frequently in nanoQSAR studies are presented.

The types of NMs covered by the nanoQSAR and grouping/read-across studies are presented schematically in Figure 5.

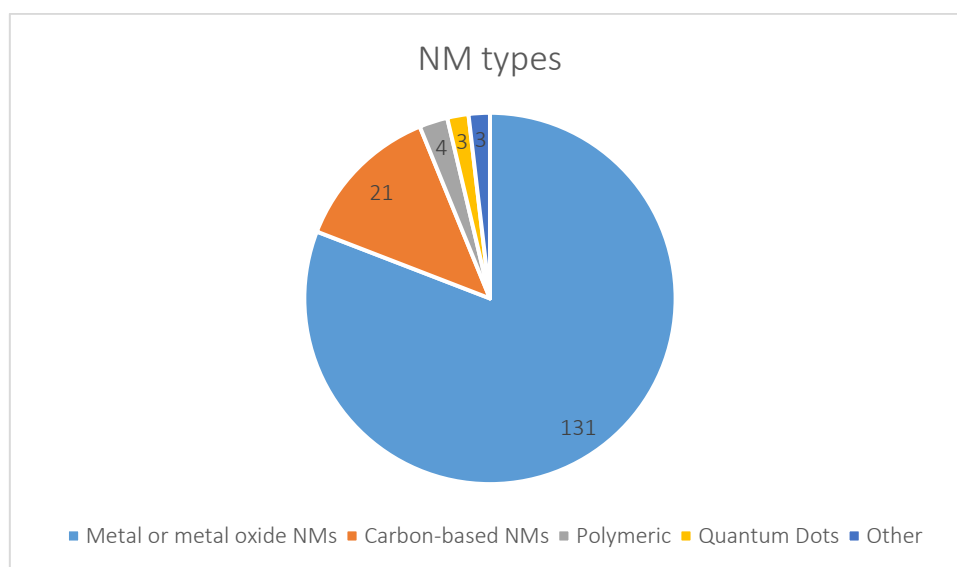


Figure 5: Types of NMs covered by the nanoQSAR and read-across studies. Some studies included models for more than one type of NMs.

The properties or toxicity behaviour that were used as endpoints for the developed nanoQSAR and read-across studies have been classified into the following categories, in order to organise and examine them collectively:

- Ecotoxicity/environmental fate: Lethality in embryonic zebrafish<sup>72</sup>, NMs concentration in plants roots<sup>73</sup>, response on Zebrafish embryo<sup>74</sup>, immobilisation of *Daphnia magna*<sup>75</sup>, etc.
- Toxicity: Cell association<sup>39</sup>, cellular uptake<sup>48</sup>, cell viability<sup>76</sup>, genotoxicity<sup>77</sup>, cytotoxicity<sup>50</sup>, oxidative stress<sup>78</sup>, exocytosis in macrophages<sup>79</sup>, etc.

- Property: zeta potential<sup>1</sup>, NMs stability<sup>80</sup>, tensile strength of graphene<sup>81</sup>, adsorption coefficients<sup>82</sup>, electronic properties<sup>83</sup>, protein corona formation<sup>84</sup>, protein carbonylation<sup>85</sup>, protein binding<sup>70</sup>, etc.

The different endpoint categories are presented schematically in respect to the number of studies in Figure 6.

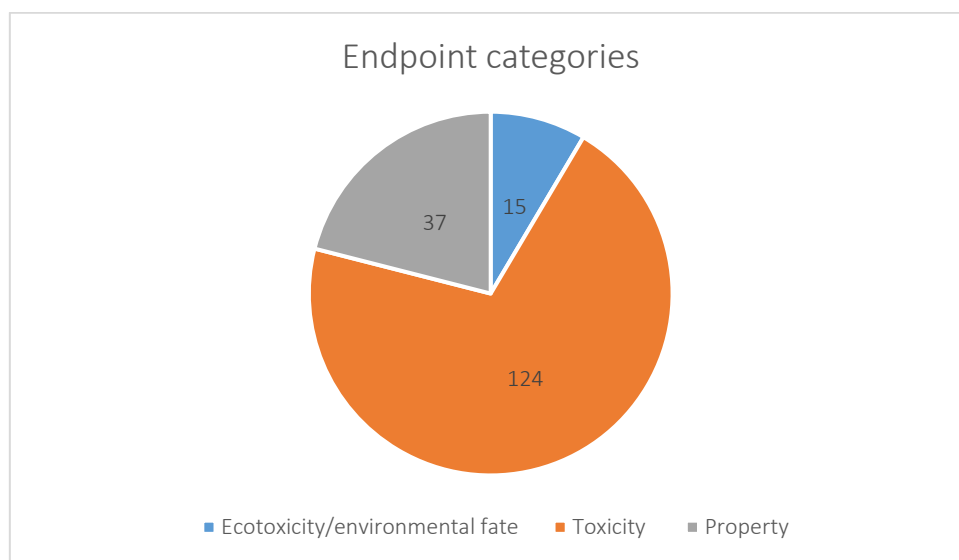


Figure 6: Endpoint types of the nanoQSAR and read-across studies. Some studies included models for more than one endpoint.

The effectiveness of a data-driven prediction method or model (such as nanoQSARs and read-across) is significantly influenced by the information encoded in the input variables. The performance of the generated predictive model may decrease, and the processing time required by the modelling algorithm may increase with the inclusion of non-informative data. The use of the datasets as is without any filtering steps may also result in overfitted models (modelling of the data errors). Variable (or feature) selection prior to modelling removes noisy attributes that do not statistically contribute to the analysis or are directly correlated to other attributes, which improves the algorithms' performance<sup>86</sup>. It also leads to the development of simpler and more comprehensive models by pointing out the variables that are significant for the studies endpoint<sup>87</sup>. Studies that included data filtering (in their majority statistical filtering such as correlation or low variance filtering) or a variable selection process prior to modelling are presented in Figure 7. In cases where variable filtering is not performed *per se* in a study, but a filtered datasets is used based on the variable selection results of former studies on the same dataset, we considered that this variable selection also covers the latter study and in the results of Figure 7 these studies are included in the "performed" class.

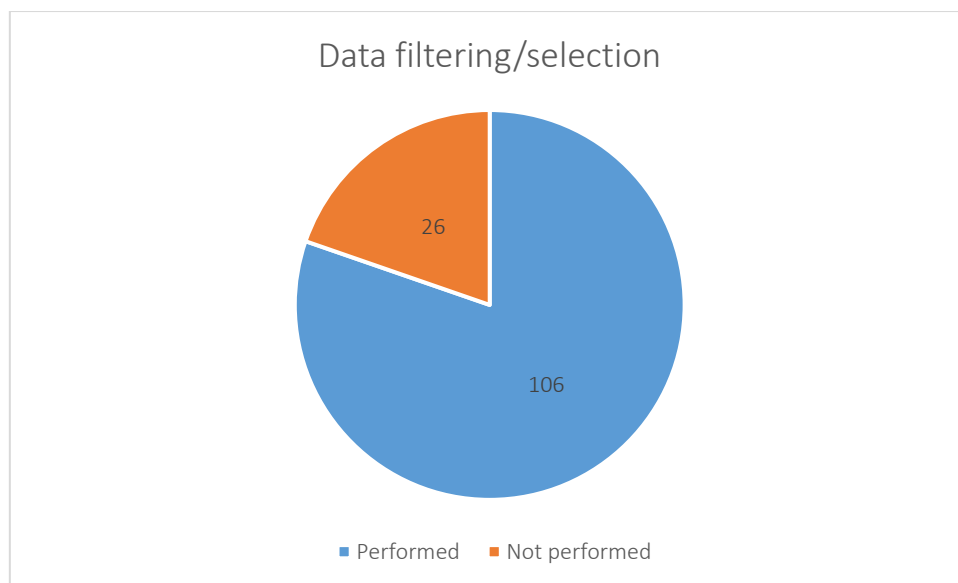


Figure 7: nanoQSAR and read-across studies including a variable selection step.

At this point we have to underline that when describing the NM data used to build the *in silico* models, the presentation of the NMs composition using only the chemical formula should not be considered sufficient when more than one phases exist. More specifically, a single material is consisted of atoms arranged in space. The types of atoms and their ratio are encoded in the material's chemical formula (e.g.,  $\text{TiO}_2$ ). However, different arrangement of these atoms in space leads to different crystal phases. In crystallography, any specific rearrangement of atoms in space is called space group. Different material phases present different extrinsic properties and thus, this information should be included in the characterisation of NMs. The calculation of molecular/theoretical descriptors for an NM structure is also impossible, in case that the NM phase is not provided. Therefore, in the field of computational simulations, to produce theoretical descriptors the material space group is needed. It is possible to access the structure of crystal materials in Crystallographic Information Files (CIF) format and their phases/space groups from different databases, e.g., the Crystallography Open database (<http://crystallography.net/cod/>). For example, if we use this database to study  $\text{TiO}_2$  NM, we can see that  $\text{TiO}_2$  forms -most commonly- two phases, rutile and anatase which correspond to the space groups of  $P4_2/mnm$  and  $I4_1/amd$  respectively (Figure 8). In many cases in this review the space groups of the NMs are not provided<sup>48,72,95–98,73,88–94</sup>.

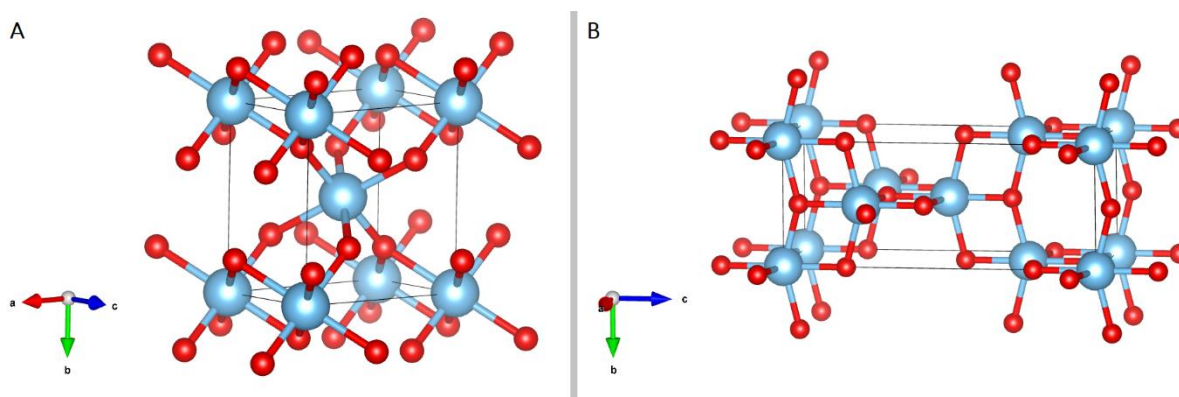


Figure 8: [A]  $\text{TiO}_2$  rutile and [B]  $\text{TiO}_2$  anatase unit cells. Titanium atoms are marked in grey and oxygen atoms are marked in red colour.



An initiative that could greatly contribute to the comprehensive description of NM structures without producing data gaps, is the development of International Chemical Identifier (InChI)-type representations for NMs (“NInChI”) that could greatly enhance nanoinformatics workflows and NMs grouping<sup>2</sup>. More precisely, the development of machine-readable text representations of the NM structure could include the encoding -in layers- of its chemical composition (considering among others the NM crystallinity), its morphology, its surface parameters, functionalisation, and ligands. To support this effort, authors also presented an alpha version of an NInChI generator user-friendly tool released via Enalos Cloud platform (<http://www.enaloscloud.novamechanics.com/nanocommons/NInChI/>, Figure 9).

The nanomaterial is being built with a layered architecture from core to outer layers

Composition	Morphology	Size (nm)	Crystal layer	Chirality layer
	Shell	t	None	n,m
Au	Sphere	d = 25.0	Au	
Ag	Shell	t = 30.0	Ag	

NInChI **NInChI=1A/Ag/msh/s30t-9/k[Fm-3m]!Ag/msh/s30t-9/k[Fm-3m]/y2&1**

Figure 9: Screenshot of the alpha version of the NInChI tool, to generate simple NM structures containing the NM’s core and potential shells and clusters.

### 5.1.2 Model development

The nanoQSAR models and the grouping/read-across strategies -to correlate the available NM descriptors to the endpoint of interest- are employing machine learning methods and algorithms. Depending on the type of endpoint (numerical value or class) different methods are employed (regression and classification methods respectively). Unsupervised clustering methodologies are also employed in grouping studies. In Table 4 the commonly used algorithms in the reported studies are presented, and in Figure 10 the number of regression and classification studies. Additionally, employing the development of novel hardware and architectures, deep learning approaches for large data analysis have also been developed for the *in silico* evaluation of NMs. Deep learning models are composed of multiple processing layers aiming to learn from representations of data with a high level of abstraction (e.g., processing and object recognition from images, audio-visual files, and speech). These models use their multilayer structure to extract complex and abstract features from the input raw data during training, thus no variable selection is needed prior to modelling. As an example, Karatzas *et al.*<sup>99</sup>, applied two deep learning methods on microscopic images of *Daphnia magna* exposed to TiO<sub>2</sub>, Ag, or AgS NMs under different media conditions. This work was done to detect regions of interest for possible abnormalities, caused by the direct or previous generations exposure to these NMs. Meanwhile, Yan *et al.*<sup>97</sup> in their study to predict physicochemical parameters and nano-bio interactions, generated features of nanostructures from photos of NPs using convolutional neural network techniques.

Table 4: Commonly used algorithms and machine learning methods or the reported nanoQSAR and grouping/tread-across approaches.

Regression Models	Classification Models	Clustering Techniques
Multiple linear regression (MLR) <sup>35,41,50,91,100,101</sup>	Decision trees <sup>49,65,85</sup>	Hierarchical clustering analysis (HCA) <sup>26,102</sup>
Partial least square regression (PLS) <sup>38,45,92,103–105</sup>	<i>k</i> -nearest neighbour <sup>36,70,75</sup>	Self-organising maps (SOM) <sup>56,93,106</sup>
<i>k</i> -nearest neighbour ( <i>k</i> NN) <sup>1,89,107</sup>	Support vector machines <sup>77,108</sup>	K-means clustering <sup>90</sup>
Multi-layer perception <sup>46,95,109</sup>	Logistic regression <sup>72,76,77,110</sup>	Principal component analysis (PCA) <sup>60,102,111–114</sup>
Neural networks <sup>41,48,73,81,115,116</sup>	Naïve Bayes classifier <sup>66,77</sup>	
Extreme gradient boosting regressor (XGB) <sup>109,117</sup>	Random forest <sup>72,77,113,118</sup>	
Random forest regressor <sup>38,41,61,90,109</sup>		

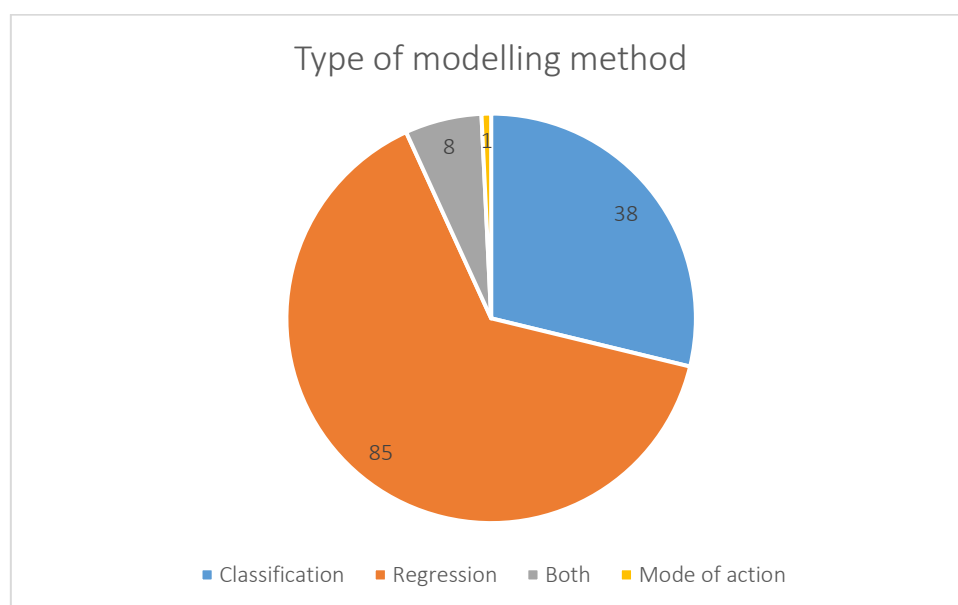


Figure 10: Type of nanoQSAR and grouping/read-across models according to the endpoint type (numerical-regression and categorical-classification). In some of the studies both regression and classification models were presented.

Considering the particularities of the nanotoxicity field (the properties of materials in nanoscale, the NMs dynamic behaviour depending on the environment, the data scarcity, the need to expedite the NMs risk assessment process), effort have been made in the development of tailor-made methods that address these needs. In the study of Rybinska-Fryca *et al.*<sup>112</sup> the concept of structure-activity prediction networks (SAPNets) is presented as a strategy that effectively correlates the description of NMs' structure with their toxicity using a system of layers made up of nodes that are equivalent to “meta-models”. Thus, the network is constructed from the endpoint to the descriptors that describe the NM structure layer-by-layer. As a result, SAPNets give enough details on what structural properties need to be changed to produce an NM with desired properties. Tsiliki *et al.*<sup>39</sup> developed an R package called RRegrs that performs optimised model selection between ten simple and complex regression algorithms, contributing in that way in the development of reliable models in the field of nanoinformatics. Varsou *et al.*<sup>34,37</sup> developed methodologies and tools to generate read-across

predictions based on different levels of similarities (thresholds) between NMs to achieve a refined selection of neighbours. The read-across predictions are based on the weighted average of neighbouring NMs thus, more similar NMs have more influence in the prediction. An advantage of these methodologies is that they are “universal”, meaning that they can be used to generate different models based on different datasets, allowing their use in different cases (endpoints, properties, or types of NMs) without any need of a grouping hypothesis. In the study of Chatterjee *et al.*<sup>59</sup> a similar workflow is presented for the refined selection of neighbours and the generation of read-across predictions based on the weighted average. Varsou *et al.* have also been focused on the development of automated and optimised workflows for the search of the best NM grouping hypothesis and the development of local MLR models through the solution of mathematical programming problems<sup>35</sup>. Finally, Gajewicz *et al.*<sup>57</sup> presented the Nano-QRA general grouping/read-across methodology that is based on the one-point-slope, the two-point formula, or the equation of a plane passing through three points, which is effective for data-gaps filling.

According to the OECD principles for the validation of QSAR models<sup>119</sup>, the QSAR models should be associated with “a mechanistic interpretation, if possible”. In this course, we recorded the studies where a discussion on the selected descriptors (descriptors space) and their influence on the endpoint is made or the studies where a mechanistic interpretation is provided for the query toxicity endpoint (Figure 11). The lack of the models’ mechanistic interpretation is possibly due to the lack of relevant knowledge (for the endpoint and the descriptors) from the part of data modellers, or due to the specific type of models. For example, in deep learning models based on images the descriptors are directly derived from them to perform classification, thus the assessment of the descriptors’ space is not possible.

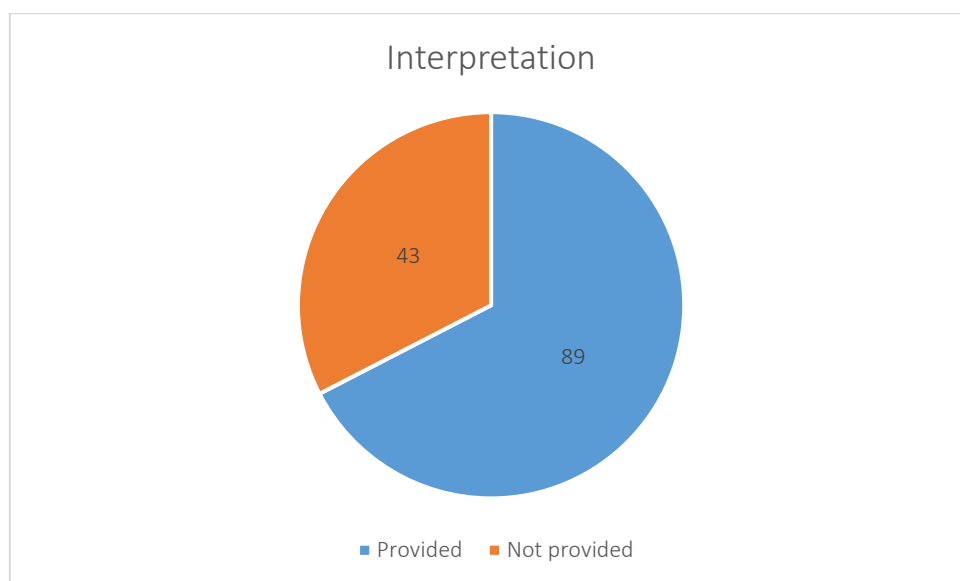


Figure 11: nanoQSAR and read-across studies, where a mechanistic interpretation of the models is provided.

Finally, another point of interest was the possibility of the models’ implementation as user-friendly tools so that they serve the broader nanocommunity. Consequently, we searched for the implementation potential of the models considering the availability of the codes and scripts via public repositories (e.g., GitHub) or directly through the relevant publications, and for simpler models (e.g., linear models) we reported whether the derived mathematical formula is presented in the publications (Figure 12). In the same figure we have included the 29 nanoQSAR and read-across models and methods that already available through a GUI (see also §5.6 and Appendix I). We can observe that the majority of models can be (or already are) implemented, thus modelers should be encouraged to make them

available as tools or -in case that they lack the time and/or the know-how to develop such tools- they should be encouraged to share them through deployment platforms such as the Jaqpot (<https://www.jaqpot.org/>) and the Enalos Cloud Platform (<http://www.enaloscloud.novamechanics.com/>).

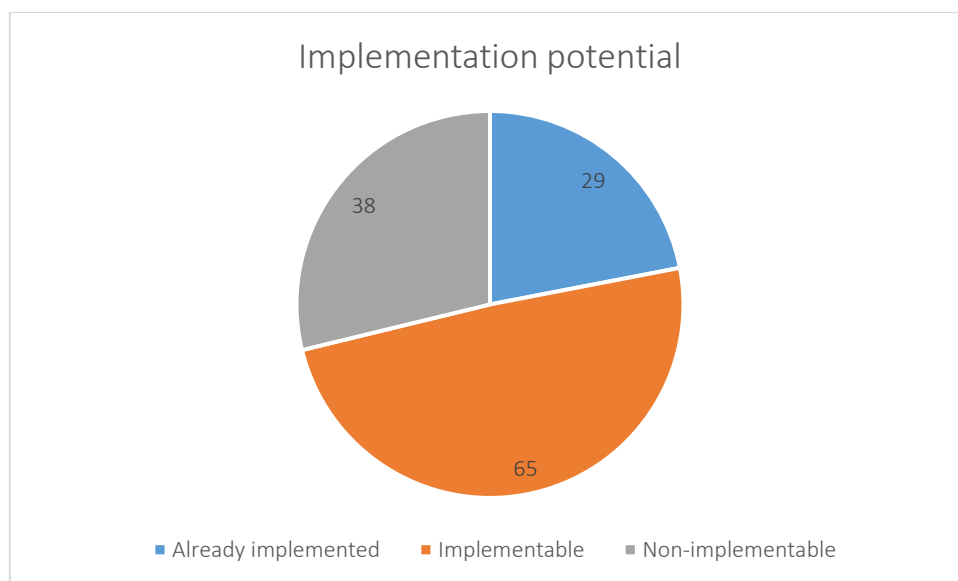


Figure 12: nanoQSAR and read-across models that already are or they can be easily implemented as user-friendly tools.

### 5.1.3 Model validation

The most important aspect in the creation and use of a predictive model is its validation, considering that it verifies the reliability of the model and its acceptability for external data. Model validation can be performed by internal and external validation techniques. In an external validation scheme, an independent data set, usually including experimentally derived values that did not participate during model development, is used to test the predictive ability of the model in “real” external data conditions. The simplest technique to perform external validation is to split the available data set into training and test sets at random<sup>41,46,61,75,120–122</sup>, commonly choosing a certain percentage of the NMs. Another broadly used algorithm to partition the data into training and test sets is the Kennard-Stone method, that allows representative selection of NM samples between sets by covering all the data space<sup>35,76,77,123</sup>. Then, a model is built based on said training data and it is applied to the remaining testing data. Finally, its capability of prediction is evaluated.

In the Literature it is highlighted that for small datasets validation scores may vary between different splits of the data<sup>124</sup>. Given that many datasets used in the field of nanoinformatics are small (less than 20 samples), an internal validation process can be applied to overcome this potential flaw<sup>86</sup>. Internal validation, which is a goodness-of-fit and robustness test<sup>119</sup> is usually performed through k-fold cross validation, a popular resampling technique. Each time, the original data set is reduced either by one compound (leave-one-out) or by a small group of compounds (leave-many-out), thus a new model is developed in every step, and the selection bias that could occur by the selection of only one training set is eliminated. Especially in the case of leave-one-out (LOO) cross validation, all the available information is exploited thus, it supports the users’ confidence in the classifier’s accuracy and it is also a deterministic process and no random splitting is involved contributing to the elimination of the selection bias<sup>86</sup>.

For reproducibility and transparency purposes, when splitting the data in a random way, authors are advised to provide the exact training and test NMs. In cases of a deterministic splitting (e.g., LOO cross

validation, Kennard-Stone method), this could be omitted. The number of studies that permit reproducibility through the availability of the training and test sets are presented in Figure 13.

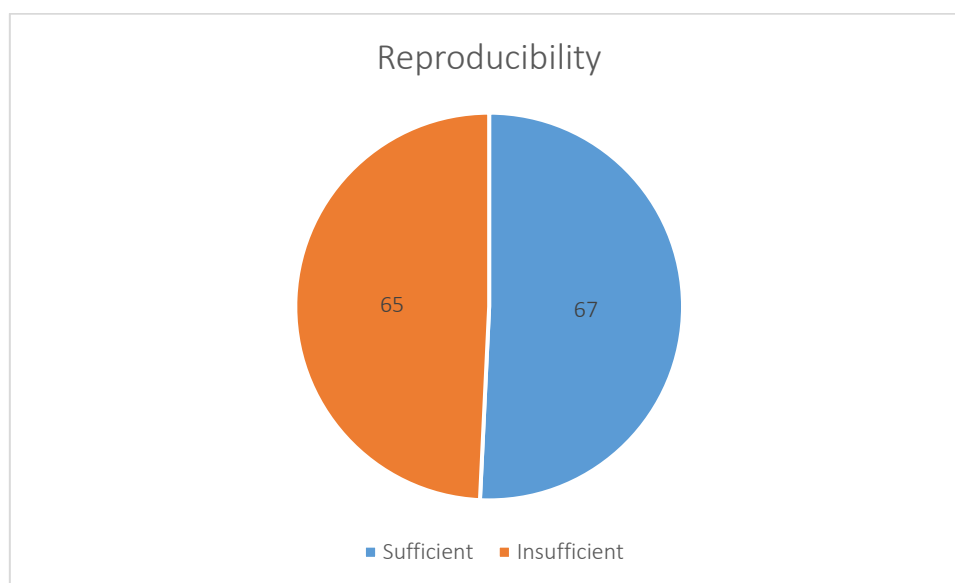


Figure 13: nanoQSAR and read-across models where training and test sets are provided, supporting the reproducibility of the methods.

The predictive model parameters are advised to be selected (whenever enough data are available) by evaluating the model both internally (e.g., following a k-fold validation scheme) and externally (e.g., using a test set that did not participate in model development). Nonetheless, in a few reported models in this review (Figure 14), the results of either validation scheme<sup>56,102,108,113,125,126</sup> are not properly presented (mainly because the grouping/read-across methods do not necessarily imply the development of a machine learning model<sup>9</sup>) or the models are trained and validated on the exact same data<sup>74,91,94,127,128</sup>.

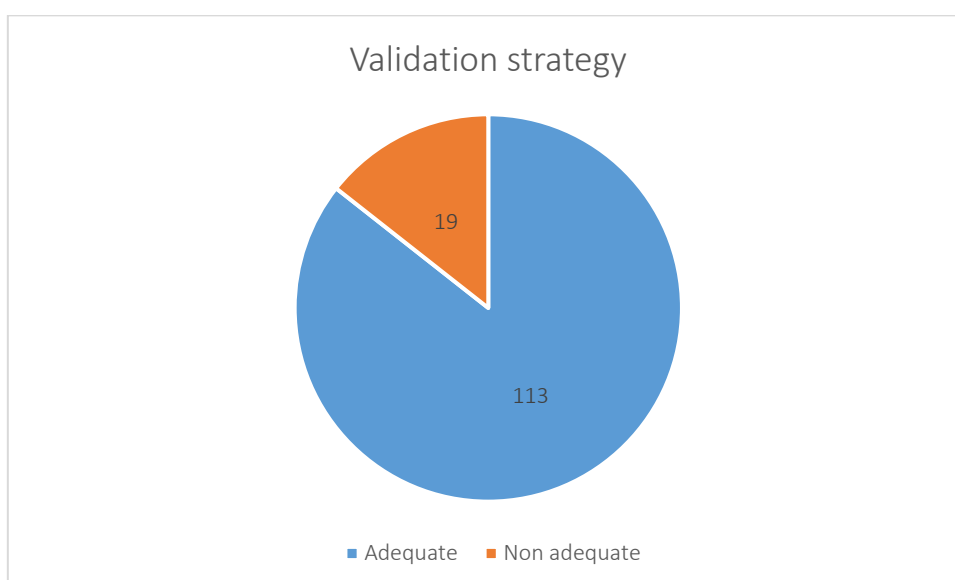


Figure 14: nanoQSAR and read-across models with an adequate validation strategy (including at least one of the internal or external validation schemes).

Moreover, to ensure the robustness of the predictive models and to investigate whether the models are based on chance correlation, the Y-randomisation (or Y-scrambling) procedure is usually

employed<sup>39,40,129,130,49,53,60,75,84,89,90,100</sup>. Within the Y-randomization scheme, the training set's endpoint values are randomly shuffled between NMs. All modelling steps are repeated several times using the original values of the independent descriptors and scrambled values for the endpoint variable. Then, when the Y-scrambled models are applied on the test NMs, if the original model is robust and reliable it is expected that the shuffled models perform poorly. Models with a provided robustness test (e.g., Y-randomization, internal cross-validation or a sensitivity analysis scheme) were recorded, and the results are presented in Figure 15.

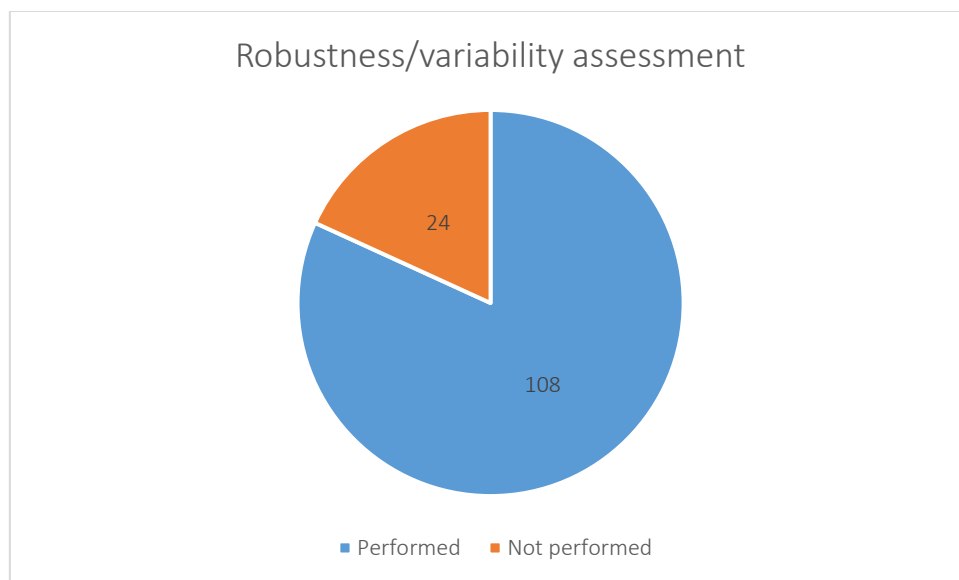


Figure 15: nanoQSAR and read-across studies where robustness tests are performed.

#### 5.1.3.1 Statistics and Metrics

Model validation is also accompanied by a number of widely used statistical measures in order to describe and assess a model's performance. Statistic metrics are chosen based on the predicted endpoint type (regression or classification methods).

As described before, methods that fall in the former category use the descriptors as independent variables and attempt to predict the value of the target feature (numerical endpoint) by estimating the relationships between the dependent and independent variables. Consequently, the model's performance is assessed based on how close the actual ( $y_i$ ) and the predicted ( $\hat{y}_i$ ) values of the endpoint are, usually by calculating the differences (residuals) between actual and predicted values. Two main categories of metrics are distinguished: the  $R^2$ -based metrics and the error-based metrics. The  $R^2$ -based metrics (such as the coefficient of determination and the external explained variance coefficient), are expected to be closer to 1 for fitter models. Error-based statistics, such as the mean squared, or absolute error are expected to be closer to 0 for more accurate models. In some of the studies other tests are employed to assess the models' performance, such as the Golbraikh and Tropsha's criterion<sup>1,36,45,53,89,129,131,132</sup>.

In Table 5 the metrics used in common base are presented.  $y_i$  and  $\hat{y}_i$  are the actual and predicted endpoint values of the  $i^{th}$  sample respectively,  $\bar{y}$  and  $\bar{\hat{y}}$  denote the average actual and predicted endpoint values respectively, and  $N$  is the total number of samples.

Different metrics are employed when a classification method is applied, since those models aim to predict the class of a target NM and not a numerical endpoint. The majority of the reviewed works use binary classification, meaning that the distinct classes are encoded as TRUE or FALSE<sup>34</sup> (other categories

may be “1” or “0”<sup>76,133</sup>, or “toxic”/ “non-toxic”<sup>75,111</sup>, etc.). The performance of these models is evaluated based on the number of correct predictions and the number of misclassifications. In studies using such methods, confusion matrices are usually presented<sup>34,99,104,111</sup>, which are essentially tables showing the number of True Positives (TP), False Positives (FP), True Negatives (TN) and False Negatives (FN). Furthermore, other classification metrics including accuracy, sensitivity (or recall or true positive rate – TPR) and specificity (or true negative rate – TNR) are widely used, while a significant number of studies calculated F1-score, Cohen’s kappa, Matthews correlation coefficient (MCC), and precision<sup>34,65,68,75,76,99,104,118,134,135</sup>. The highest value most of those statistics may acquire, demonstrating a perfect prediction, is 1. However, Cohen’s kappa and MCC can get negative values that show no agreement between predictions and actual data: 1 demonstrates a perfect prediction/agreement between actual and predicted values, 0 demonstrates a random prediction, and -1 demonstrates a false prediction/total disagreement<sup>136</sup>. Another indication of performance is the receiver operating characteristic (ROC) curve, which illustrates the diagnostic ability of the classifier as its discrimination threshold is varied<sup>118,121</sup>. More specifically, a ROC curve plots TPR on a vertical axis and false positive rate (FPR) of a horizontal axis and later the area under the curve is calculated which is expected to be closer to 1 when performance is good.

Table 5 and

Table 6 present the most common quantitative validation measures found in the reviewed studies, for regression and classification models respectively.

Table 5: Most common quantitative measures of goodness-of-fit and predictivity for regression models. These metrics are presented in detail in Appendix VI.

Metric	Formula	Definition <sup>136</sup>
Coefficient of Determination ( $R^2$ )	$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad [1]$	Estimation of the variation of predictions that is explained by the regression.
Pearson's Correlation Coefficient	$r = \frac{\sum_{i=1}^N (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^N (y_i - \bar{y})^2 \sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}})^2}} \quad [2]$	Measure of the strength of association between variables.
Mean Absolute Error (MAE)	$MAE = \frac{\sum_{i=1}^N  y_i - \hat{y}_i }{N} \quad [3]$	Measures the average of the absolute error values.
Mean Squared Error (MSE)	$MSE = \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N} \quad [4]$	Measures the average of the squares of errors.
Root Mean Squared Error (RMSE)	$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}} \quad [5]$	Root square of average squared error.
External explained variance coefficient ( $Q_{\text{ext}}^2$ ) or external validation coefficient ( $Q_{F1}^2$ )	$Q_{\text{ext}}^2 \text{ or } Q_{F1}^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_{\text{train}})^2} \quad [6]$	The sum of squares on the numerator runs over the test NMs and the reference total sum of squares on the denominator is calculated comparing the predicted response of the test NMs with the average endpoint of the training set <sup>119</sup> .
External validation coefficient ( $Q_{F2}^2$ )	$Q_{F2}^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_{\text{test}})^2} \quad [7]$	The sum of squares on the numerator runs over the test NMs and the reference total sum of squares on the denominator is calculated comparing the predicted response of the test NMs with the average



		endpoint of the test set <sup>59</sup> .
Cross validated correlation coefficient ( $Q_{cv}^2$ ) <sup>50</sup>	$Q_{cv}^2 = 1 - \frac{\sum_{i=1}^N (y_i - \widehat{y}_i^{cv})^2}{\sum_{i=1}^N (y_i - \bar{y}_{train})^2} \quad [8]$	Authors applied the internal validation (cross validation leave-one-out technique) to reduce probability of the model's overfitting to the training data, and to measure robustness of the model on the presence/absence of particular NMs in the training set <sup>50</sup> .
Squared leave-one-out cross-validation correlation coefficient ( $Q_{abs}^2$ ) <sup>137</sup>	$Q_{abs}^2 = 1 - \frac{\sum_{i=1}^N (y_i - \widehat{y}_i^{LOO})^2}{\sum_{i=1}^N (y_i - \bar{y}_{train})^2} \quad [9]$	This metric is calculated for the training set NMs based on the leave-one-out predictions <sup>137</sup> .

One source of confusion (that can also impede the direct comparison of studies on the same data) is the use of different quantitative measures for the validation of regression models under the same symbol. A source that many researchers consult is the OECD validation principles for QSAR models<sup>119</sup> but still not all employed measures are covered by this guide and also, different definitions exist in the Literature for the same statistic metric and/or symbol. A characteristic example of misinterpretations is the use of symbols  $r^2$  and/or  $R^2$  or of the term “R-squared” to describe the “squared Pearson correlation coefficient” (Eq. 10) or the “coefficient of determination respectively” (Eq.1).

$$r^2 \text{ or } R^2 = \left( \frac{\sum_{i=1}^N (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^N (y_i - \bar{y})^2 \sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}})^2}} \right)^2 \quad [10]$$

Where  $N$ , is the number of samples,  $y_i$  and  $\hat{y}_i$ , are the actual and predicted endpoint values of the  $i^{\text{th}}$  sample respectively, and  $\bar{y}$  and  $\bar{\hat{y}}$ , are the average endpoint values of the experimental and predicted values respectively.

In some studies, authors use the “R-squared” term without providing the explicit formula (or at least a definition)<sup>73,105,138–141</sup> or they use the types and the definitions interchangeably<sup>50,90,100,142</sup>. This may potentially lead to erroneous comparisons regarding the reliability of models trained on the same data and could be overcome with the presentation of the “R-squared” formula on the results, instead of the mere use of the name or of the symbol. For example, in the publication of Pan *et al.*<sup>54</sup> authors use the squared correlation coefficient (Eq. 10) in model validation. Later, they compare their results to the model of Puzyn *et al.*<sup>50</sup> which use the correlation of determination in their model validation (Eq. 1). In fact, this misunderstanding is justified because Puzyn *et al.* use the term “squared regression

coefficient” in the main document of the publication, however they provide the R-squared formula in the respective supplementary material, and this corresponds to the “coefficient of determination”.

We present some additional examples of possible misunderstanding: for example, Puzyn *et al.*<sup>50</sup> and Fourches *et al.*<sup>137</sup> use different symbols and definitions to describe the same metric (see  $Q_{\text{abs}}^2$  and  $Q_{\text{CV}}^2$  in Table 5). In addition, the OECD definition of external explained variance ( $Q_{\text{ext}}^2$ )<sup>119</sup> is the same to the external validation coefficient ( $Q_{\text{F1}}^2$ )<sup>59</sup>. In cases where the definitions are provided along with the publication (even as supplementary material) it is possible to make meaningful comparisons between models, otherwise misconceptions could arise.

Table 6: Most common quantitative measures of goodness-of-fit and predictivity for classification models.

Metric	Formula	Definition <sup>136</sup>
Sensitivity or Recall (True Positive Rate)	$TPR = \frac{TP}{TP + FN} \quad [11]$	The probability of a positive case conditioned on truly being positive.
Specificity (True Negative Rate)	$TNR = \frac{TN}{TN + FP} \quad [12]$	The probability of a negative case conditioned on truly being negative.
Fall-out (False Positive Rate)	$FPR = \frac{FP}{FP + TN} \quad [13]$	The proportion of false positives among all negative cases.
Accuracy (ACC)	$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad [14]$	The proportion of correct predictions in the total number of samples.
Precision (Positive Predictive Value)	$PPV = \frac{TP}{TP + FP} \quad [15]$	The proportion of true positive among all positive observations.
F1-score	$F1 = \frac{2TP}{2TP + FP + FN} \quad [16]$	Harmonic mean of sensitivity and precision.
Matthews Correlation Coefficient (MCC)	$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad [17]$	Measure of association for the two binary variables, takes values between [-1,1].
Cohen's kappa	$\kappa = \frac{2(TP \times TN - FP \times FN)}{(TP + FP)(FP + TN) + (TP + FN)(TN + FN)} \quad [18]$	Measure of inter-rater reliability, takes values between [-1,1].

The selection of the appropriate measures of accuracy when quantifying a model's performance and fitness is a matter of the model's type (e.g., while neural networks aim to minimize squared error or cross entropy, support vector machines and boosting are designed to maximize accuracy)<sup>143</sup>. In addition, other issues should be considered, e.g., RMSE-type metrics may be susceptible to outliers (the squaring of prediction errors may produce bias in the final mean error value, considering that higher error values are more influential than lower prediction errors)<sup>136,144</sup>. The RMSE (squared) units may also be considered "unnatural"<sup>136</sup>.  $R^2$  or  $Q^2$ -type metrics may be considered more intuitive, however the data range and the distribution of the endpoint values around their mean value in the train/test set may impede the reflection of the actual prediction errors<sup>144</sup>. Another example is related to the popularity of F1-score as a measure of accuracy. F1-score has received criticism due to its "conceptual flaws"<sup>145</sup>. More specifically, the F1-score lacks symmetry and varies when the positive class is switched with the negative class. In addition, the F1-score value is independent of the true negatives and thus, it fails to fully capture the binary classifier's performance<sup>146</sup>. MCC and Cohen's kappa are suggested as alternatives to F1-score<sup>145,146</sup>. To sum up, the choice of one performance measure instead of others leads to various compromises as no metric can be ideally employed for every modelling application<sup>143</sup>. To better evaluate the quality-of-fit and the predictivity of a model, it is advisable to present various metrics to acquire a complete and thorough validation of the prediction performance. In this course, when evaluating the reviewed models in this project, we applied a cut-off value of four metrics as the acceptable lower limit to consider the reported models properly evaluated (see §5.1.6). The studies (models and tools) that included a sufficient number of statistic measures (at least 4) is presented in Figure 16.

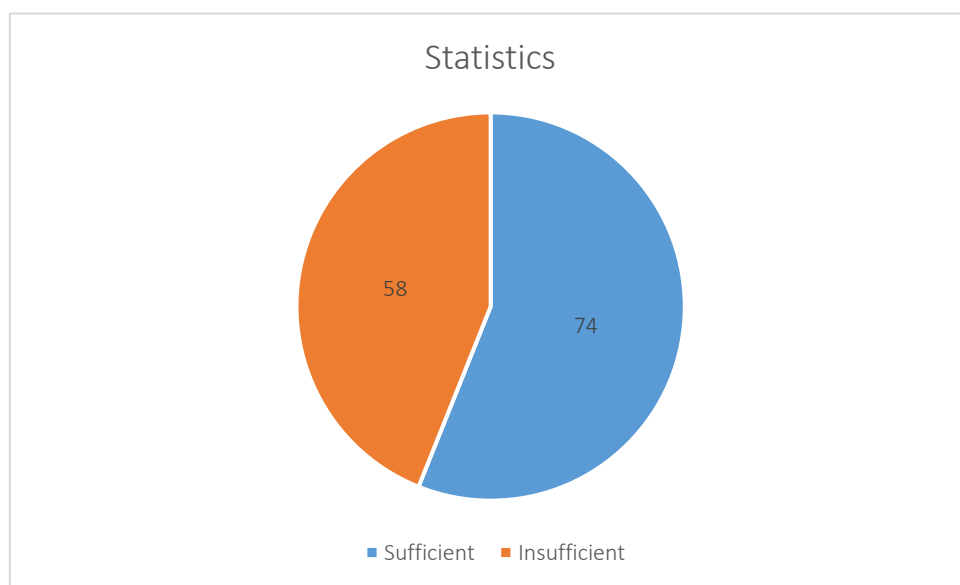


Figure 16: nanoQSAR and read-across studies reporting more than 4 statistical metrics.

#### 5.1.4 Applicability domain definition

The reductionist nature of nanoQSAR models associates them with limitations in terms of chemical structures, physicochemical properties and mechanisms within a response space<sup>119</sup>. In general, extrapolated predictions (inference outside of the known training dataset) are considered less accurate than interpolated ones (inference between known training NMs). The definition of the DoA determines the area of reliable predictions based on the training set and it is necessary for describing the limitations of the model. The degree of similarity between the NM of interest and the model training set is determined by different approaches, such as the leverage approach and a distance-based method. The applicability domain is ordinarily defined after model validation and assesses whether the predicted

values of the test or an external set are reliable or unreliable, thus strengthening the user's confidence in the results. The DoA boundaries may vary depending on the desired trade-off between the models' extent of use and the reliability of their predictions<sup>119</sup>.

The most common method for DoA definition, the leverage method<sup>33,45,53,55,76,77,142</sup>, provides a measure of the distance of the compound from the centroid of its training set. Based on the extend of extrapolation, the leverage  $h$  for each compound is calculated from the diagonal elements of the Hat matrix<sup>147,148</sup>:

$$H = X(X^T X)^{-1} X^T \quad [19]$$

The warning leverage ( $h^*$ ) is the critical value for a reliable prediction, which means that if a compound's leverage  $h$  is larger than  $h^*$  ( $h > h^*$ ), the prediction is considered unreliable:

$$h^* = \frac{3p}{N} \quad [20]$$

Where  $X$ , is the table containing the training data,  $p$ , is the number of descriptors used in model<sup>76,77,148</sup> (or it is the number of descriptors used in model plus one<sup>33,119,149,150</sup>), and  $N$ , is the number of samples in the training set.

The leverage approach can be visualised via a Williams plot<sup>41,60,151–154</sup>, which interpolates the standardised cross-validated residuals versus leverage values for each compound of the training set<sup>119,150</sup>.

Moreover, another approach for defining the DoA is to use distance-based (Euclidean/Manhattan/Mahalanobis) methods<sup>107,130</sup>. Those methods consider similarity measurements based on the distances among training and test compounds compared to the predefined applicability domain threshold. Different modelers determine different thresholds for distance-based techniques because there are no clear guidelines in the field<sup>148</sup>. A common threshold calculation<sup>88,133,155</sup> is presented here:

$$AD = \langle d \rangle + Z\sigma \quad [21]$$

Where  $\langle d \rangle$  and  $\sigma$  are the average and the standard deviation of the distances respectively, and  $Z$  is an empirical parameter value whose default value is 0.5. In case that the distance from an external compound to its nearest neighbour is larger than the threshold  $AD$  then the prediction is labelled as unreliable.

In other studies<sup>44,70,89</sup>, the DoA threshold was calculated in a slightly different way, based again on the previous equation (Eq. 21). In this case, all Euclidean distances between all training samples, as well as the mean value of these distances, are determined first. Next, the new average value ( $\langle d \rangle$ ) and standard deviation ( $\sigma$ ) of the distances contained in the subset of training samples -with shorter distance than the average distance of all training samples- are calculated.

Further applicability domain definition approaches found in the reviewed studies include the standardised residuals deviation<sup>130</sup> and the probability density distribution-based method<sup>156</sup>. The former is a degree-of-fit method, and the latter identifies a highest density region from the total probability mass. Other methods are also reported such as, the training descriptors range-based domains<sup>35,46,72,122,139</sup> (where untested NM descriptors' values should be within the ranges defined by the training NMs in order to consider a prediction as "reliable"), the use of the standardization approach<sup>140,157</sup> developed by Roy *et al.*<sup>158</sup>, and the multiple threshold method<sup>159</sup>.

Finally, in some studies the applicability domain is not defined by a specific method. Some examples are presented below:

- In the study of Forest *et al.*<sup>104</sup> the applicability domain is considered as the area of (hydr)oxide NPs.
- In the study of Aschberger *et al.*<sup>102</sup> the applicability domain is defined from the size and physicochemical parameters of the training MWCNTs (e.g., longer or surface modified MWCNTs or MWCNTs with higher content of oxidising impurities are outside the applicability domain).

In most cases assessed in this review, the definition of the applicability domain was performed as part of the standard modelling workflow. However, a remarkable number of studies did not present the applicability limits of the respective models questioning the level of reliability of the produced predictions (Figure 17).

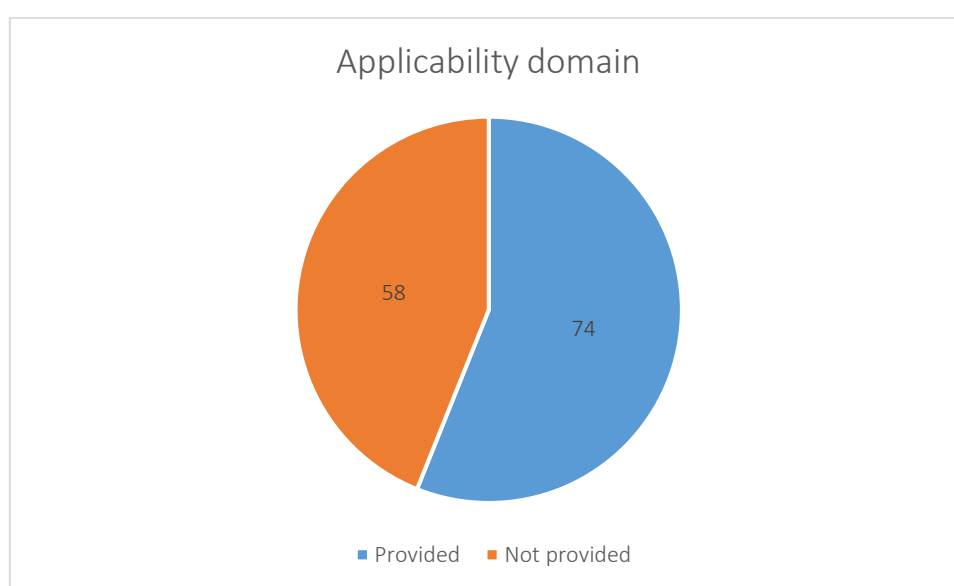


Figure 17: nanoQSARs and read-across models and tools with a defined DoA.

#### 5.1.5 Model dissemination

A harmonised, established documentation methodology with a commonly accepted terminology can be particularly helpful when sharing research results, given the complexity of each individual approach and wide range of methods, cases and implementations. This need is addressed by the development of standardised reporting formats that can be used as a reporting and evaluation system for computational models' predictions that aims to offer industry and regulators reliable and FAIR data. Within this study, it was investigated whether the included *in silico* approaches have been accompanied with a documentation (through the tool interface or in supplementary files of scientific publications) using a specific template (e.g., the QSAR Model Reporting Format-QMRF or the Modelling Data-MODA generalisation reporting template). These reports were initially developed for QSAR models of chemical substances or for material simulations, but they can be used without major modifications for nanoinformatics approaches.

The MODA reporting template is developed by the European Materials Modelling Council (EMMC)<sup>160</sup>. MODA is a detailed template meant to guide users towards a complete high-level documentation of material models by providing all the necessary aspects for description, reproducibility, curation, and interfacing with other models. A simple consecutive workflow with data-based stand-alone model can be seen in Figure 18:

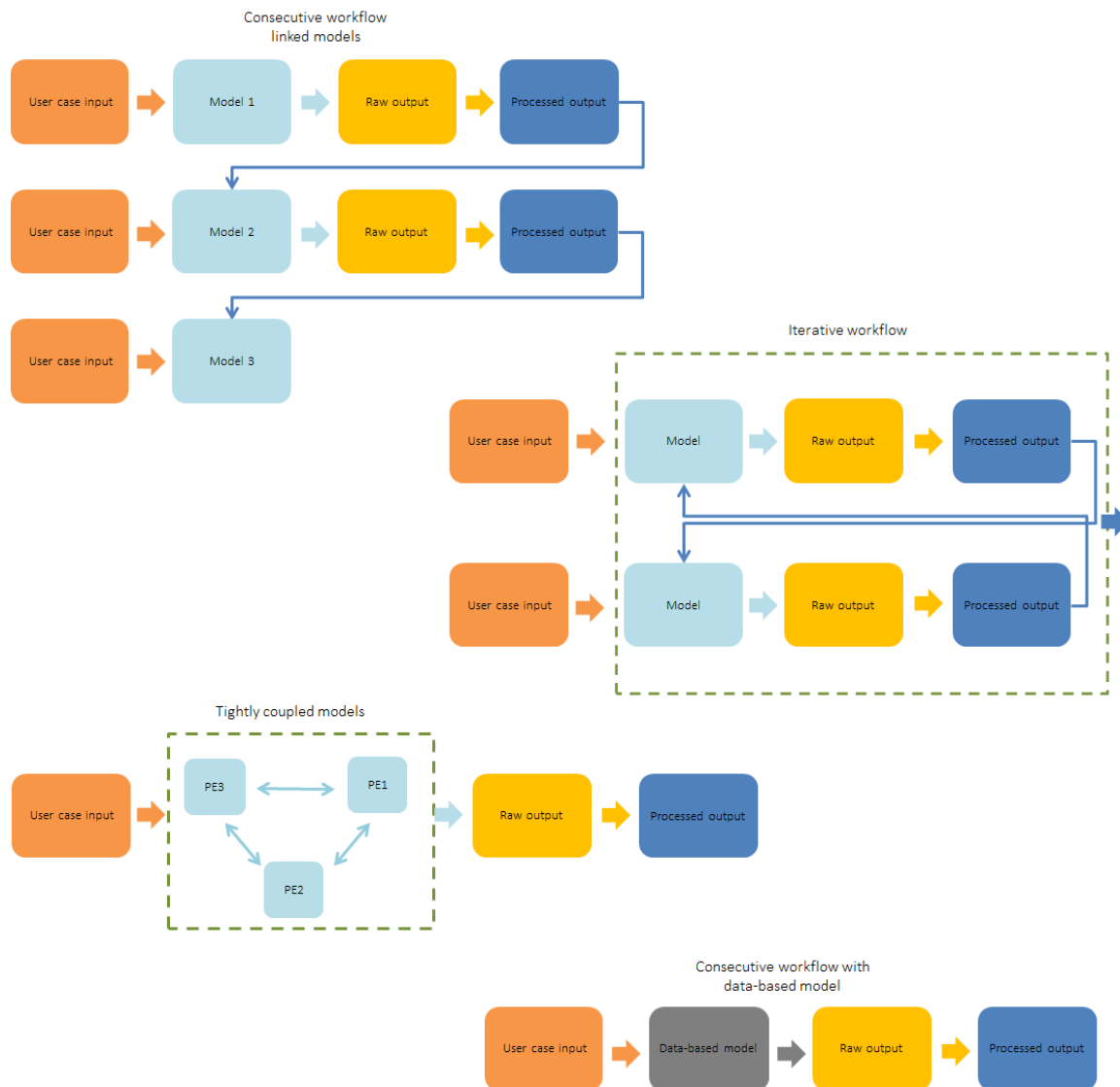


Figure 18: A graphical workflow scheme that should be used to document materials modelling according to MODA depending on the type of reported models. Image adapted from [https://emmc.eu/wp-content/uploads/2021/05/MODA\\_Workflow\\_templates\\_2021.pptx](https://emmc.eu/wp-content/uploads/2021/05/MODA_Workflow_templates_2021.pptx)

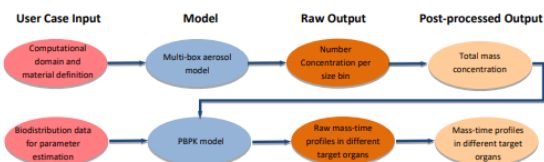
A MODA template consists of an introductory chapter for the simulation overview and three or four chapters according to the models' nature. During the overview of the simulation, one must provide a general description of the user case without giving away any modelling information, the chain of models and access conditions for the software or database. The first chapter of the report refers to the aspect of the user case/system to be simulated, determining the material, geometry, time lapse and manufacturing process. When a physics-based model is reported, the second chapter indicates the generic physics of the model equation (e.g., entity, materials relations, simulated input) and the third chapter provides information on the solver and computational translation of the specifications (numerical solver and parameters, software tool, time step, computational representation, and boundary conditions). After that, during the post-processing chapter, the output is discussed, along with the methodologies and the margin of error. In case that a data-based model is reported, a sole chapter indicates the database type and the generated equation.

An example of a model report using MODA template is presented in Figure 19.

#### 4. MODA DOCUMENTS

##### MODA #1 – An integrated multi-box aerosol and TiO<sub>2</sub> PBPK model

OVERVIEW of the simulation		
1	USER CASE	Coupling of external and internal exposure tools for predicting the deposited dose of TiO <sub>2</sub> nanoparticles in the human respiratory system and the subsequent biodistribution to other internal organs after occupational exposure.
2	CHAIN OF MODELS	<b>MODEL 1</b> Multi-box aerosol model for generating the spatiotemporal evolution of a selected nanoparticle concentration in interior spaces. <b>MODEL 2</b> PBPK model for describing the distribution of TiO <sub>2</sub> nanoparticles in humans after inhalation exposure.
3	PUBLICATION ON THE SIMULATION	Tsiros, P., Chetaniou, N., Tsoumanis, A., Jensen, A.C.O., Jensen, K.A., Melagraki, G., Lynch, I., H. Sarinvisit, H., Afantitis, A. Towards an in silico Integrated Approach for Testing and Assessment of nanomaterials: from predicted indoor air concentrations to lung dose and biodistribution. <i>Environ Sci Nano</i> 2021 (In review)
4	ACCESS CONDITIONS	The models have been developed in the NanoSolveIT research infrastructure project under grant agreement n° 814572 and can be freely accessed in <a href="https://explorepubk.cloud.nanosolveit.eu">https://explorepubk.cloud.nanosolveit.eu</a>



##### MODEL 1: Multi-box aerosol model

1 ASPECT OF THE USER CASE/SYSTEM TO BE SIMULATED	
1.1	<b>ASPECT OF THE USER CASE TO BE SIMULATED AND HOW IT FORMS A PART OF THE TOTAL USER CASE</b> A multi-box spatio-temporal model of nanomaterial (NM) exposure in a closed environment (room). The model represents a closed room divided into multiple areas (boxes) with each box having its own properties based on the type of the box. The model is driven by NM emissions data, and predicts spatio-temporal environmental concentrations in each box based on the size distribution of the NM particles.
1.2	<b>MATERIAL</b> carbon black, TiO <sub>2</sub> , TiO <sub>2</sub> /AgX, user defined materials
1.3	<b>GEOMETRY</b> Currently assumes spherical NMs.
1.4	<b>TIME LAPSE</b> Varies from seconds to days depending on the time horizon that the user has specified.
1.5	<b>MANUFACTURING PROCESS OR IN-SERVICE CONDITIONS</b> Not applicable.
1.6	<b>PUBLICATION ON THIS ONE SIMULATION</b> A. C. O. Jensen, M. Dal Masso, A. J. Koivisto, E. Belut, A. Meyer-Plath, M. Van Tongeren, A. Sánchez Jiménez, I. Tuimman, M. Domat, J. Toftum and I. K. Koponen, Comparison of Geometrical Layouts for a Multi-Box Aerosol Model from a Single-Chamber Dispersion Study, <i>Environments</i> , DOI:10.3390/environments5050052.
2 GENERIC PHYSICS OF THE MODEL EQUATION	
2.0	<b>MODEL TYPE AND NAME</b> Continuum Model.
2.1	<b>MODEL ENTITY</b> Mass population balances.

2.2	<b>MODEL PHYSICS/CHEMISTRY EQUATION PE'S</b>	<b>Equations</b> The model includes a number of physiological and substance-specific parameters which are reported in detail in Section 2.1 of the manuscript, Eq. 1-6. <b>Physical quantities for each equation</b> The model includes a number of physiological and substance-specific parameters which are reported in detail in Section 2.1 of the manuscript.
<b>MATERIALS RELATIONS</b>		<b>MR Equations</b> The model includes a number of MR which are reported in detail in Section 2.1, Eq. 1-6. <b>Physical quantities/descriptors for each MR</b> The model includes a number of MR which are reported in detail in Section 2.1, Eq. 1-6.
2.4	<b>SIMULATED INPUT</b>	Not applicable.

3 SPECIFIC COMPUTATIONAL MODELLING METADATA	
3.1	<b>NUMERICAL SOLVER</b> Dormand Prince integrator
3.2	<b>SOFTWARE TOOL</b> Java.
3.3	<b>TIME STEP</b> The time step is automatically set by the solver.
3.4	<b>COMPUTATIONAL REPRESENTATION</b> Not applicable
3.5	<b>PHYSICS EQUATION, MATERIAL RELATIONS, MATERIAL BOUNDARY CONDITIONS</b> Not applicable
3.6	<b>ADDITIONAL SOLVER PARAMETERS</b> The relative and absolute tolerance are both set to 10 <sup>-6</sup> .

4 POST PROCESSING	
4.1	<b>THE PROCESSED OUTPUT IS CALCULATED FOR</b> The raw output is the NM number concentration-time profile in each bin in cm <sup>-3</sup> . The final output is the average NM concentration-time profile over all bins in µg/m <sup>3</sup> .
4.2	<b>METHODOLOGIES</b> Not applicable.
4.3	<b>MARGIN OF ERROR</b> The model does not include uncertainty quantification.


Figure 19: Example of a MODA report for the multi-box aerosol model<sup>161</sup>.

Other modelling approaches included in this review used the QMRF reporting template provided by the Joint Research Centre (JRC) of the European Commission. The QMRF templates are considered “appropriate documentation for summarising and reporting key information on QSAR models, including the validation results of the studies. The information included in these reports is structured according to the OECD principles for the validation of QSAR models<sup>162–164</sup>:

1. A defined endpoint, intending to ensure transparency on the predicted endpoint, usually referring to a physicochemical property, a biological effect or an environmental parameter related to the structure that can be measured (expressed as a dependent variable) and modelled.
2. An unambiguous algorithm, aiming to ensure clarity in the description of the model's algorithm, which is basically the form of the relationship between the descriptors and the endpoint. In this point, the descriptor dataset (divided in training and testing sets) and data generation, the type of the mathematical model (e.g., machine learning), and the model's parameters and values are expected to be presented.
3. A defined DoA, describing the limitations of the model and assessing whether the predicted values are reliable or unreliable. The DoA definition uses a variety of approaches (distance-based, leverage, etc.) aiming to determine the degree of similarity between the chemical of interest and the model training set.
4. Appropriate measures of goodness-of-fit, robustness and predictivity, which are in fact statistical parameters describing how the model performs internally and externally. Such statistics include the square correlation coefficient ( $R^2$ ), RMSE, MAE for regression models and confusion matrices, chi-square ( $\chi^2$ ), accuracy, sensitivity for classification models.
5. A mechanistic interpretation, if possible, in order to associate the descriptor and the endpoint from a scientific viewpoint.



An example of a QMRF report for a nanoQSAR model is presented in Figure 20.

	<b>QMRF identifier (JRC Inventory):</b> To be entered by JRC
	<b>QMRF Title:</b> Tree-based regression model for predicting the zeta potential of nanomaterials (NMs) in water
	<b>Printing Date:</b> 17-Jun-2021

<p><b>1.QSAR identifier</b></p> <p><b>1.1.QSAR identifier (title):</b> Tree-based regression model for predicting the zeta potential of nanomaterials (NMs) in water</p> <p><b>1.2.Other related models:</b> None</p> <p><b>1.3.Software coding the model:</b> R version R64 4.0.3 Open-source Language and Environment for Statistical Computing <a href="http://www.R-project.org">http://www.R-project.org</a></p>	<p>Zeta potential in water.</p> <p><b>3.6.Experimental protocol:</b> None</p> <p><b>3.7.Endpoint data quality and variability:</b> Both endpoint data and species data were extracted from the online database: <a href="https://db.nanopharos.eu/Queries/Datasets_zul">https://db.nanopharos.eu/Queries/Datasets_zul</a></p>
---	---

<p><b>2.General information</b></p> <p><b>2.1.Date of QMRF:</b> 15 June 2021</p> <p><b>2.2.QMRF author(s) and contact details:</b> Chrysoula Gousiadiou National Technical University of Athens (NTUA) cgousiadiou@gmail.com</p> <p><b>2.3.Date of QMRF update(s):</b> N/A</p> <p><b>2.4.QMRF update(s):</b> N/A</p> <p><b>2.5.Model developer(s) and contact details:</b> Chrysoula Gousiadiou National Technical University of Athens (NTUA) cgousiadiou@gmail.com</p> <p><b>2.6.Date of model development and/or publication:</b> 15 June 2021</p> <p><b>2.7.Reference(s) to main scientific papers and/or software package:</b></p> <p><b>2.8.Availability of information about the model:</b> None</p> <p><b>2.9.Availability of another QMRF for exactly the same model:</b> None</p>	<p><b>4.Defining the algorithm - OECD Principle 2</b></p> <p><b>4.1.Type of model:</b> Extreme Gradient Boosting</p> <p><b>4.2.Explicit algorithm:</b> XGBoost Trees XGBoostTree algorithm (Boosting Trees method) is an ensemble learner, creating a final model based on a collection of weak individual models (decision trees). In XGBoostTree the individual models are not built on completely random subsets of data and features, but sequentially by putting more weight on instances with wrong predictions and high errors. In each round of training, the weak learner is built, and its predictions are compared to the observed output values. The difference between predicted and true values represents the error rate of the model. The gradient is used to minimize a loss function and serves to find the direction in which to change the model parameters in order to (maximally) reduce the error in the next round of training by "descending the gradient" (gradient descent).</p> <p><b>4.3.Descriptors in the model:</b> [1]Coating categorical physicochemical [2]Type of coating (Anionic, Cationic, Neutral) categorical physicochemical [3]Geometric Surface Area nm2 physicochemical [4]Corresponding sphere diameter nm physicochemical [5]Molecular Weight g/mol molecular [6]Number of Oxygens - molecular [7]Pauling electronegativity - molecular</p> <p><b>4.4.Descriptor selection:</b> The initial dataset contained 23 features. For the modelling we used 7 independent features after performing a backwards selection of descriptors (Recursive Feature Elimination, RFE) with Random Forest algorithm (randomForest package - version 4.6-14). Random Forest has a built-in feature selection as well as variable importance estimation utilized for the RFE approach.</p> <p><b>4.5.Algorithm and descriptor generation:</b> N/A</p> <p><b>4.6.Software name and version for descriptor generation:</b></p> <p><b>4.7.Chemicals/Descriptors ratio:</b> 0.14</p>
---	---

<p><b>3.Defining the endpoint - OECD Principle 1</b></p> <p><b>3.1.Species:</b> Nanoparticles</p> <p><b>3.2.Endpoint:</b> [1][2]P-CHEM 4.29. Zeta potential</p> <p><b>3.3.Comment on endpoint:</b> Zeta potential in water (continuous).</p> <p><b>3.4.Endpoint units:</b> mV</p> <p><b>3.5.Dependent variable:</b></p>	<p><b>5.Defining the applicability domain - OECD Principle 3</b></p> <p><b>5.1.Description of the applicability domain of the model:</b> Leverage Threshold: 0.42</p>
---	---

Figure 20: Example of a QMRF report (first 2 pages) for a nanoQSAR model (<https://zetapot.cloud.nanosolveit.eu/>) that predicts the NMs zeta potential in water.

In this review, the models found in Literature or repositories supplied with a MODA or QMRF report are depicted schematically in Figure 21. As it can be observed, the majority of the models are not accompanied with a standardised report. Therefore, impediments may be presented when stakeholders aim to reproduce or use an *in silico* methodology of interest. One of them is the time that should be spent on reading the respective scientific publication or to contact the modellers, with the aim to derive the modelling parameters-workflow. Later, stakeholders need to build the same model and validate it in terms of reproducibility, which may require a trial-and-error process till the exact same model is developed. This may impede stakeholders to use and consequently profit from the developed *in silico* approaches.

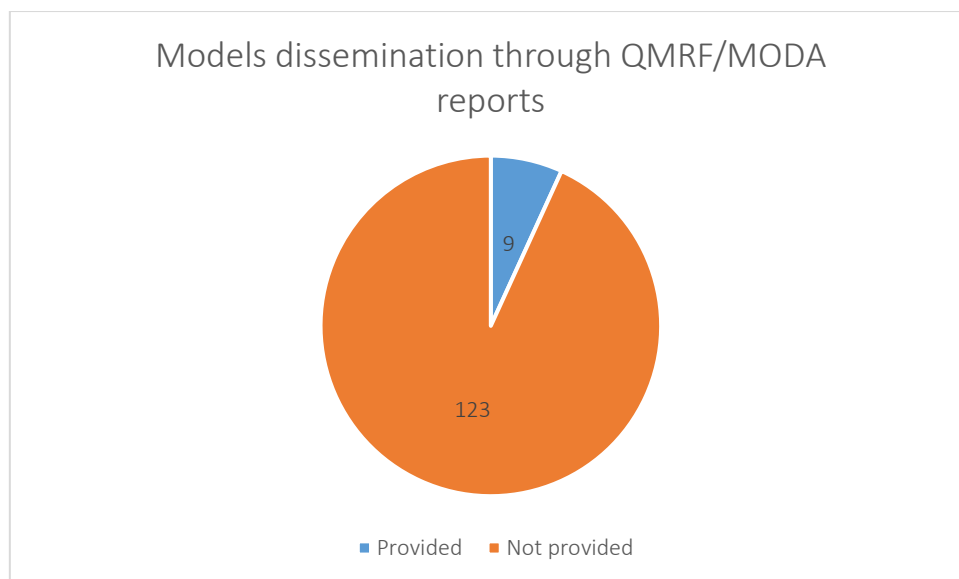


Figure 21: nanoQSAR and read-across models supplied with a standardised report.

### 5.1.6 Quality assessment of the nanoQSAR and read-across studies

After the analysis of the collected studies, a quality score was assigned to each one of them based on the extracted information. This overview process was emphasised solely on nanoQSAR and read-across studies, provided that they presented sufficient information for a detailed analysis. Exposure and causal models, PBPKs and AOPs have substantial differences in the development and validation compared to ML data-driven methods (e.g., the confidence in an AOP is assessed by WoE approaches<sup>165</sup> and not with conventional ML validation methods), which did not allow detailed quality assessment under the criteria presented here. The studies related to these methods are presented separately in the next sections (§5.2 to 5.5). Based on the key aspects defined above, a series of criteria were devised:

1. Availability of sufficient information to enable evaluation of the validity and suitability of the *in silico* validated alternative methods (e.g., are the train test sets provided? are all modelling steps and parameters included to reproduce the model?).
2. Filtering of the descriptors or variable selection performed prior to the model development (e.g., to remove noisy attributes and avoid overfitting phenomena). If variable selection is performed in previous publications, we considered that this criterion is met for the current tool/model.
3. Information on the validation strategy (e.g., if at least one of internal/external validation is performed) and if the endpoint of the alternative method is well-defined and fits the model.
4. Report of sufficient (at least 4) validation metrics-statistics to evaluate the model's performance.
5. Information on the alternative method applicability, boundaries, and gaps.
6. Provided robustness or variability tests (e.g., internal validation, Y-randomization test) or sensitivity analysis of the developed methodology.
7. Information on the mechanistic interpretation of the model or descriptors' space assessment.
8. Implementation potential of the models: If the model is already available (e.g., a GUI is already provided for the model), and or if there are possibilities of development user-friendly web applications based on the reported models (e.g., if the code or the equations are available through the publication).
9. Information on the development and reproducibility of the alternative method, e.g., QMRF, MODA reports.
10. Information on alternative methods' data and meta-data availability (e.g., in supporting information files, in public databases).

Later, for each study and for each criterion a value of “1” was recorded in case that the relevant criterion was satisfied, and “0” in case that the criterion was not met. Based on the total number of satisfied criteria, a score value was attached. The possible scores are:

- More than eight (8) criteria reported: 5 points,
- Six (6) to seven (7) criteria reported: 4 points,
- Four (4) to five (5) criteria reported: 3 points,
- Two (2) to three (3) criteria reported: 2 points,
- Zero (0) to one (1) criterion reported: 0 points,
- Any extra defined criterion reported (e.g., availability of a GUI via a web-service or a stand-alone application for the developed model or methodology, scripts availability on public repositories, development of general methods independent from the data, optimization and/or automation of the computational workflow, use of the ECHA read-across workflow, etc.): +1 point. It is noted that as in the criterion no. 8 both already implemented and implementable models are considered, an extra point is given here for the models accessible via GUIs.

The nanoQSAR and read-across methodologies in this review were evaluated according to the previous criteria and the scores are depicted graphically in Figure 22, Figure 23, and in Figure 24. For comparison purposes in Figure 22, the extra given points (if any) at each methodology are depicted in a different colour. Considering the number of studies that satisfy at least 4<sup>b</sup> common criteria (82%) and at least 6 criteria (47%), we can conclude that the quality of the assessed models and of their reporting are quite high. A few future improvements in the model development and/or reporting can be made to satisfy more quality criteria. Moreover, 38% of the assessed studies included at least one extra criterion. The extra criteria that were satisfied are grouped and depicted schematically in **Error! Reference source not found.** It is clear that there is a tendency toward developing user-friendly tools and publishing the scripts in public repositories, which should be encouraged in the future to help with the creation of FAIRer models. Another positive aspect that should be also supported in the future is the integration of the optimisation and automation procedures in some of the assessed studies and the development of data-independent methodologies (general methodologies that can be used with different data and produce different models). Finally, it is noted that only 3 out of the 63 assessed studies published post 2017 (when the first version of the relevant guidance was published<sup>12</sup>) were based on the ECHA workflow for grouping and read-across. The general character of this workflow and the absence of practical information on assays and thresholds seems to have prevented its use in more studies.

It should be noted at this point, that a universal and at the same time totally objective score cannot encompass the characteristics of all the reviewed studies and tools. As an example, for deep learning models, variable selection is considered unmeaningful, as the descriptors are derived directly from images<sup>166</sup>. Another example is the study of Serra *et al.*<sup>167</sup> (also presented in §7.2.2.4) where the applicability domain is not explicitly defined, as the goal of their work is to visualise phenotypic entities under a grouping framework and not to predict a specific endpoint. This does not imply that these studies are inferior to other models solely based on the score value. In addition, authors may not include all the modelling aspects (e.g., sensitivity analysis, all robustness tests) on their publications due to “space restrictions” or time limitations to report their models, even if these steps are performed in their workflows and scripts. Therefore, the reported score values should not be considered as a strict grading system but encode rather a general trend in the reported models. The development of an objective, as much as possible, quality assessment schema for *in silico* methodologies should be one of the future steps that relevant research should focus.

---

<sup>b</sup> In their majority the models that fulfilled 4-5 criteria and are scored with 3, are not supplied with a standardised report. We consider that apart from this criterion, the quality of such models is fair.

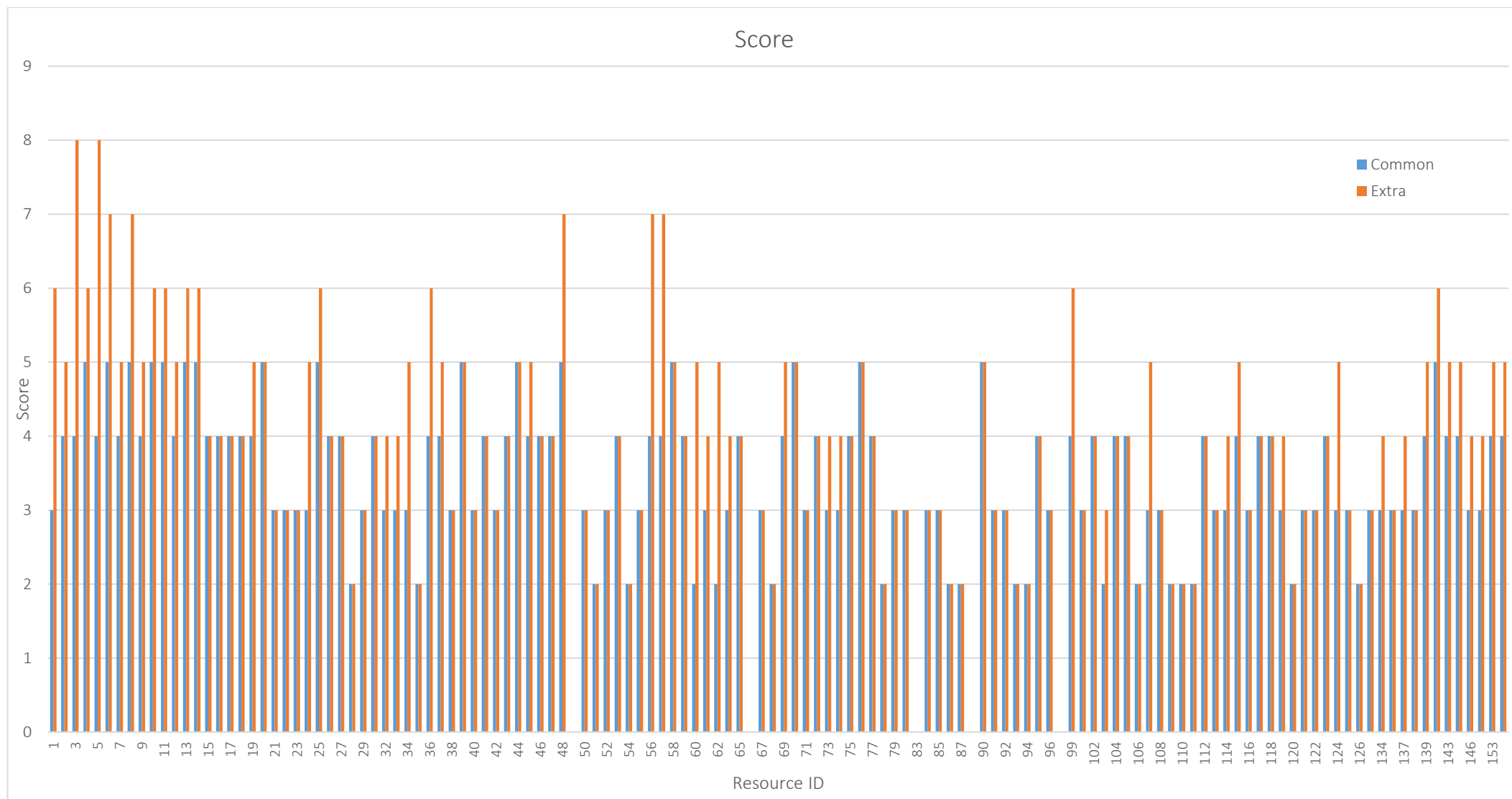


Figure 22: Score of nanoQSAR and read-across studies based on the mentioned criteria. The maximum score considering the 10 criteria is 5 and it is common for all studies (blue bars). Any other criterion is marked separately, and the final score is presented in orange colour. The IDs on X axis correspond to the studies presented in Appendix I.

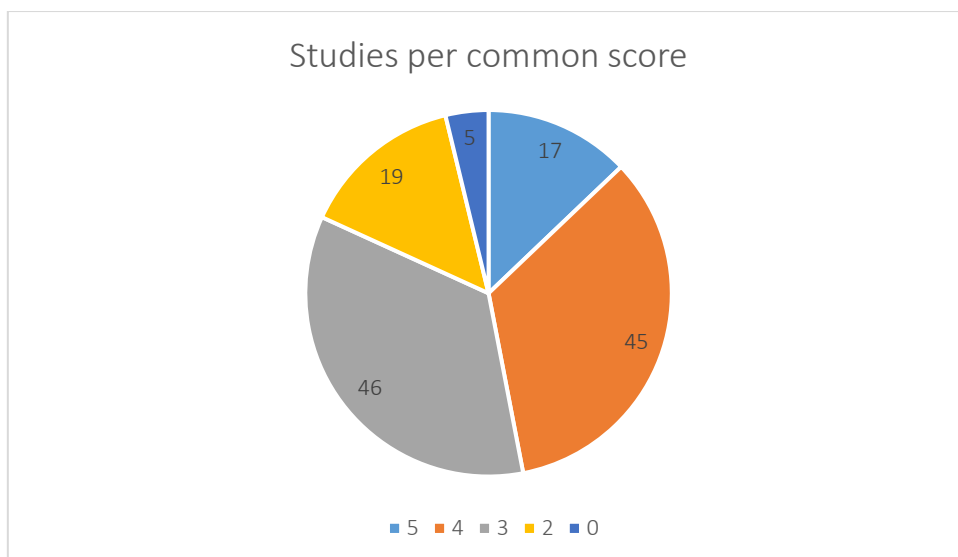


Figure 23: Number of nanoQSAR and read-across studies in respect to the common score values (0, 2, 3, 4, 5) based on the quality assessment criteria.

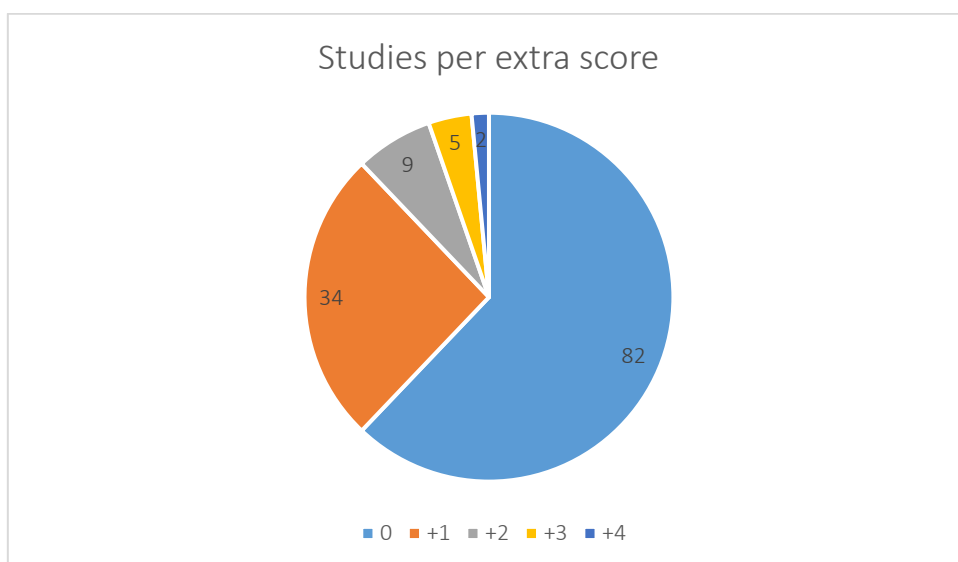


Figure 24: Number of nanoQSAR and read-across studies in respect to the extra score values (0, 1, 2, 3, 4) based on the quality assessment criteria.

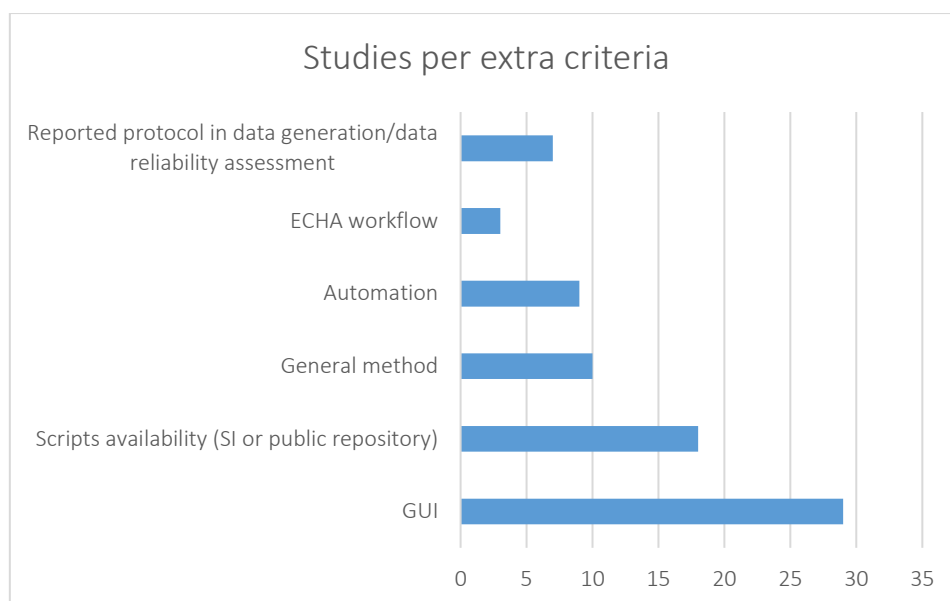


Figure 25: Number of assessed studies satisfying extra criteria apart of the 10 quality control criteria.

## 5.2 Physiologically-based pharmacokinetics models

Physiologically-based pharmacokinetics (PBPK) models are used as a tool for NM risk assessment on human and animal health, since they translate external exposures into toxicologically relevant internal doses on target organs within organisms. Eleven published PBPK approaches aiming at concentration estimations of different types of NMs (Figure 26) in independent body compartments were assessed in the present work.

Such models in general do not allow variable selection, yet in the majority of the reviewed papers emphasis was placed on the variability assessment of the physiological parameters, commonly obtained from literature. Methodologies such as sensitivity analyses were primarily applied for either the total or the most influential parameters used (depending on the studied system these can be the phagocytizing cells NPs uptake capacity<sup>168</sup>, the diffusion and partition coefficients of the kidneys and body compartments<sup>169</sup>, etc.), with the exception of Liang *et al.*<sup>13</sup> who provided a sensitivity analysis for all mice organs. Apart from the necessary physiological parameters needed for the models' development, supplemental information included data obtained *in vivo*, NM's physicochemical parameters and experiment conditions such as exposure times. The main modelling approach of those papers is the use of mass balance/mass transfer equations, simulating the compartmental biodistribution as a system of differential/algebraic equations. Kinetic models and first-order processes were included in some cases.

Tsiros *et al.*<sup>15</sup> published an integrated computational framework for assessing NM biodistribution in the human respiratory system, after several exposure scenarios in an occupational setting (indoor environment). An external exposure model, calculating the NM concentrations as a function of time, is connected with internal biokinetics exposure models for acute and chronic timescales. In this work, a lung exposure model is compared to a more realistic PBPK model: the former calculated the accumulated NM mass in the alveolar, tracheobronchial, and head airways regions, while the latter computes the same results in many regions of the human body, making it more suitable for longer-term exposure calculations.

Titanium levels for each organ in mice were estimated by Bachler *et al.*<sup>14</sup>, within the application domain from 15 to 150 nm, according to their ability to cross the capillary wall of the organs and to be

phagocytosed in the mononuclear phagocyte system. Additionally, an attempt was made in determining the disposal mechanism of ingested TiO<sub>2</sub> NPs. Lin *et al.*<sup>170</sup> developed a PBPK model for predicting the amounts of gold NPs in different tissues, blood, and urine, within an application size domain of 13–100 nm. Animal-to-human extrapolation of NM pharmacokinetics was performed from mice, rats, and pigs, respectively, to humans.

Li *et al.*<sup>169</sup> developed such a model to examine the biodistribution of poly(lactic-co-glycolic) acid formulations with different polyethylene glycol content. Here, multivariate regression analysis was also performed to build the relationship between nanoparticle properties and biodistribution parameters. Similarly, Li *et al.*<sup>168</sup> used the PBPK approach to calculate biodistribution of intravenous exposure of polyethylene on glycol-coated polyacrylamide (PAA-peg) NPs in rats.

Zazo *et al.*<sup>171</sup> simulated stavudine (antiretroviral drug) biodistribution after administering a 40 nm Au NP-based drug delivery system in rats. Another model developed by Pery *et al.*<sup>172</sup> used pharmacokinetics to predict the absorption and distribution of technetium-labelled carbon nanoparticles (Technegas) in 24 body compartments, while parameters were estimated using Bayesian-Markov chain Monte Carlo techniques. Chen *et al.*<sup>173</sup> presented a PBPK model to trace zinc oxide and zinc nitrate in mice by calculating the partition coefficients for ZnO and Zn(NO<sub>3</sub>)<sub>2</sub>, excretion or elimination rates over time.

Lastly, opposed to other papers that constructed models for specific NMs, Sahneh *et al.*<sup>174</sup> developed a PBPK model for a broad NM range, whose goal is to predict the impact of biocorona formation kinetics on interspecies extrapolation of NP-biodistribution based on basal metabolic rate (BMR) and blood circulation time. They provided an internal validation by comparing rates between two species, concluding that longer circulation times cause different responses between rodents and humans in reaching target cells.

Apart from Sahneh *et al.*<sup>174</sup>, the greater part of reviewed papers validated the developed models via external validation, by comparing their results to other *in vivo* studies or by overlaying the simulated concentration-time profiles with observed data. Most papers did not provide sufficient metrics and statistics, given the limited ability of the PBPK methodology to be validated, therefore the statistical significance of the models could not be evaluated. Last but not least, experimental studies on the behaviour of NMs in the body should be used to support the development of robust and reliable PBPK models and increase their confidence.

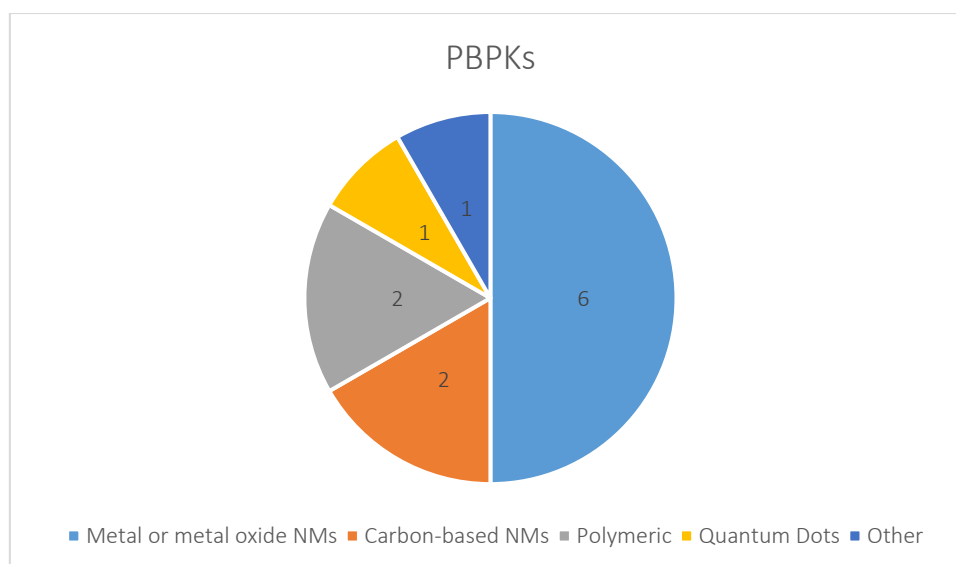


Figure 26: Types of NMs covered by the PBPK studies.

### 5.3 Molecular dynamics simulations

Molecular dynamics (MD) simulations are used as a means of evaluating the potential risk assessment of NMs computationally, as an alternative to experimental approaches. In fact, MD simulations can be considered as virtual experiments with an extra benefit: The gained knowledge on the nano-bio interactions using MD (e.g., underlying mechanism) cannot be easily obtained experimentally. As these virtual experiments are designed from scratch, there is a possibility of simulating systems that are not extensively studied experimentally (e.g., graphene in biological environments). By performing simulation methodologies, significant information can be extracted about NM's interaction with biological molecules, cells, or lipid membranes. In fact, researchers can understand the mechanism of the interaction of different types of NMs with biological systems in molecular level and they can later predict macroscopic NM properties. Therefore, the possible toxicological effects induced by the interaction between NMs and biomolecules can be predicted and integrated in the SbD of novel NMs.

This review included 28 molecular dynamics simulations, examining possible underlying mechanisms of NM cytotoxicity. As mentioned above several tools, such as NAMD, LAMMPS, and GROMACS, were utilised by the vast majority of the studies to perform MD simulations, while they were either all-atom or CG models.

Is it observed (Figure 27) that most of the reviewed MD studies have developed simulations for graphene or graphene oxide (GO) NMs (graphene nanosheets and flakes) owing to the material's ability to penetrate into the biological membrane and extract large amounts of phospholipids. However, in order to reduce the computational time required for simulating a biological organism and a graphene sheet (GS), a bilayer lipid membrane is used instead of a complex cell membrane. Tu *et al.*<sup>175</sup> investigated GS interactions with *E. Coli* inner and outer membranes using all-atom MD and performed quantum mechanics calculations on the oxidation pathways. Both types of molecular mechanisms for the graphene-induced degradation of *E. Coli* cell membranes, namely the severe insertion/cutting and the destructive lipid extraction, suggest that GS can induce serious membrane stress, and thus significantly reduce cell viability. The same system was studied by Duan *et al.*<sup>176</sup> (2015) using the Lerb-Klinowski force field for GO nanosheets, concluding that Bovine serum albumin (BSA)-coated GO sheets are less cytotoxic than the uncoated ones. The protein corona protects the cell membrane from being penetrated due to the reduction of available surface interacting with the lipid bilayer and due to obstacles imposed by the BSA geometry.



Lipid membrane (phosphatidylcholine bilayer) perturbation from graphene nanosheets was studied intensively. Li *et al.*<sup>177</sup> determined that small pristine in the biological environment spontaneously penetrate the lipid bilayer starting from corners and asperities that are more likely to overcome the energy barrier at room temperature. Following that, they observed spreading and complete penetration driven by hydrophobic forces between the graphene and the bilayer core. Similarly, Zhu *et al.*<sup>178</sup> assumed that GSs may affect the flip-flop transition of membrane lipids responsible for signalling events and regulating the morphology of the cell. Also, GSs could dramatically affect the distribution of lipid, which is presented by the average density distribution of lipid molecules in the plane of the bilayer. MD simulations by Dallavelle *et al.*<sup>179</sup> interpreted differences in the GS navigation to the membrane according to their size and their functionalisation patterns. Specifically, it was revealed that edge functionalised GOs always lie on top of the lipid bilayer, while randomly functionalised ones penetrate the membrane when their size is substantially greater than that of GS that only adhere to it, thus improving flake penetration. Li *et al.*<sup>180</sup> showed that small nanosheets sink into lipid membranes without changing their structures, but they could dramatically affect the fluidity of nearby lipids. As the size of nanosheets increases, they could cause membrane corrugation and even severe deformation of the overall liposome structure.

Additionally, several studies investigated the lipid membrane of eukaryotic cells (Dipalmitoyl-phosphatidylcholine bilayer) perturbation from GS. One of them was conducted by Puigpelat *et al.*<sup>181</sup> and showed that graphene penetration starts at corners or asperities and tends to migrate to regions where cholesterol molecules are sparse. Likewise, Chen *et al.*<sup>182</sup> revealed that GO sheets behave differently than pristine graphene nanosheets, due to the hydrophilic GO surface arising from the oxygen-contained groups. On the contrary, GSs diffuse easily into the bilayer with its basal plane parallel to lipid tails, due to hydrophobic interaction between graphene and lipid tails. Chong *et al.*<sup>183</sup>, who performed MD simulations for blood proteins concluded that the protein adsorption is mainly enthalpically driven through strong  $\pi$ - $\pi$  stacking interactions between GO and aromatic protein residues, in addition to hydrophobic interactions. Feng *et al.*<sup>184</sup> showed that GOs can cut through the hydrophobic interface of a HIV-1 integrase homo-dimer, inducing the dissociation of the dimer by favourable hydrophobic interaction.

Apart from graphene NMs, a few studies focused on carbon nanotubes (CNTs). Firstly, according to Zhang *et al.*<sup>127</sup>, the interaction affinity between the carbon-based NPs (CNPs) and the SARS-CoV-2 RNA fragment increased in the order of fullerenes < graphenes < CNTs. A study by Zhu *et al.*<sup>185</sup> determined that stiff nanotubes beyond a critical length are compressed by lysosomal membranes causing persistent tip contact with the inner membrane leaflet, eventually leading to cell death. A few more studies worth mentioning are that of Melby *et al.*<sup>186</sup> and Lehn and Katz<sup>187</sup> involving gold NPs. The former indicated that proteins may present high affinity sites for NP binding, which can result in NP accumulation, while the latter showed that amphiphilic NPs can insert into the low-curvature face of a lipid ribbon, representative of insertion into defect-free lipid bilayers, if they first come into contact with solvent-exposed lipid tail protrusions.

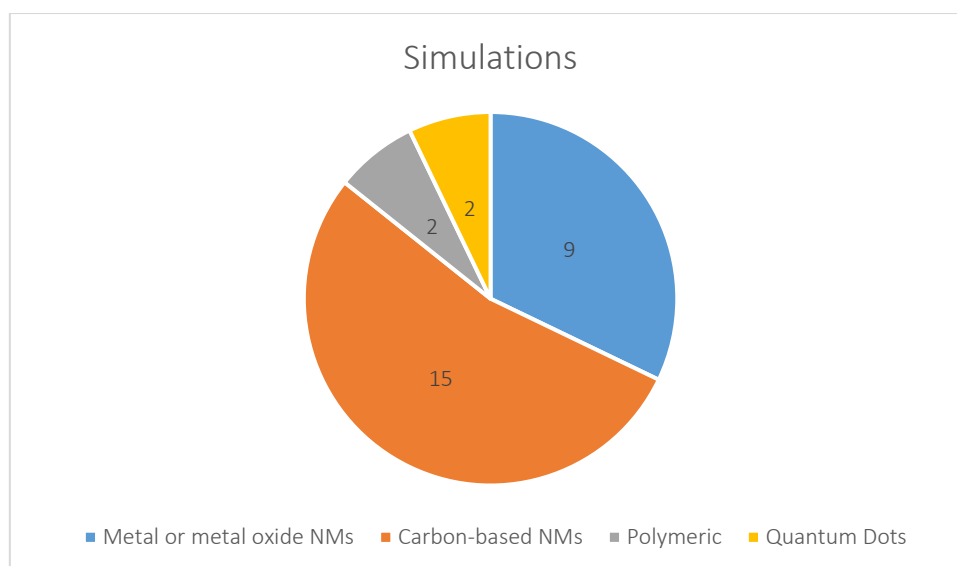


Figure 27: Types of NMs covered by the simulation studies.

Finally, in the reviewed MD studies different types of interactions of NMs with biomolecules (and mainly lipid membranes) were assessed. These systems of lipid bilayers-water-NM are used as a simplified version of a biological system comprising a complex cell membrane, a rich extracellular environment, and the NM structure. The study of interactions of cell-membranes and NMs are of particular interest for the understanding of the NMs cytotoxicity in living organisms. For instance, many of the MD studies listed above revealed that graphene sheets can enter the biological membrane and extract significant amounts of phospholipids. As a result, holes may be formed in the membrane and disturb its integrity. Such conclusions are valuable when studying the cell viability and cytotoxicity of NMs and can be included in the NMs risk assessment processes.

#### 5.4 Adverse outcome pathways

As previously stated, AOPs are sequential chains of events and mechanisms, which detect molecular alterations happening within cells and organs. Several studies are focusing on AOPs, analysing their potential links to the presence of NMs, and using them as tools which enable the detection of toxicity effects in organisms. They present AOPs as toxicity predictive workflows that combine existing knowledge of NM's toxicokinetic-toxicodynamic properties and known biological responses and pathways, trying to improve risk assessment of NMs for human health and environmental safety. It is noted that in AOP studies the DoA cannot be defined due to a lack of knowledge regarding how the size-associated features of NMs affect the AOP transition from the molecular initiating event to an ultimate AO.<sup>188</sup>

Labib *et al.*<sup>189</sup> developed a linear, simplified (hypothetical) AOP associated solely with MWCNTs-induced lung fibrosis (using cellular sensing of the substance as the MIE), based on transcriptomics data derived *in vivo* and demonstrated how such data can be used to generate pathway-based points of departure (POD). This work utilises gene expression profiles (transcriptomics data) from NM-exposed mice that developed pulmonary fibrosis and physicochemical properties of MWCNTs and identifies similar omics-based perturbed pathways. Then, benchmark doses (BMDs) of apical endpoints were calculated for each pathway and transcriptomics BMDs were derived for each CNT, where a time-dependent trend in the toxicity pathways was observed. The AO of the AOP claims that excessive extracellular matrix deposition results in the development of lung fibrotic lesions, while the results indicate that transcriptomic data can be used as an alternative to epidemiological data.

A different approach was proposed by Vietti *et al.*<sup>190</sup>, aiming to obtain simple and fast test methods for predicting the potential health hazard of CNTs. In this work, an overview of current knowledge and gaps on KEs involved in lung fibrosis development by CNTs was presented, with the intention of drafting an AOP. A large number of KEs was described based exclusively on CNT-specific literature, and their relationship to lung fibrosis (LF) was assessed. That led to a complex network of known pathways, showing the dynamics between macro-molecular interactions (fibroblasts, macrophages, and epithelial cells), along with the various possibilities of developing LF induced by high aspect ratio CNTs-inhalation, aspiration, or injection. Also, they observed a correlation between specific physicochemical properties, such as length and diameter, and CNT's fibro-genic activity. Finally, this work suggests that studying direct and indirect CNT's pro-fibrotic activities can be a risk assessment tool, providing help in developing prevention strategies related to NM-induced diseases.

Instead of focusing entirely on literature for a specific NM, Gerloff *et al.*<sup>23</sup> combined it with known chemically-induced mechanistic toxicological processes. They presented how an AOP associated with chemical-induced toxicity can aid in describing the pathogenesis of health hazards caused by any NM, provided that the downstream key events remain the same. In this work, the example of liver toxicity induced by chemicals and NMs was used, since the liver is one of the main organs where inhaled NPs can usually be found. During the case study of liver fibrosis, protein alkylation leading to LF (AOP ID: 38) was compared and merged with lysosomal uptake-induced LF (AOP ID: 144) and it was clear that the NM-induced AO shares similarities with the pathway caused by chemicals especially on the major downstream KEs. A challenge of this methodology is how to differentiate generic key events with NM-specific KEs, as well as the lack of MIEs explaining the initial interaction between NMs and molecules/biological systems, as NMs could firstly cause mechanical or physical damage to tissues and organs. More specifically, a single NM could in some cases initiate multiple interactions at the same time and in other cases they could act through non-specific interactions, causing toxicity via second messengers. On the contrary, the respective MIEs for chemicals are described adequately, thus it was proposed that NM-induced AOPs should use common KEs rather than MIEs.

A data-driven systems toxicology workflow was developed by Grafström *et al.*<sup>191</sup>, combining the AOP approach with omics-based methods in order to screen and classify the effects of engineered NMs. A cell-based, HTS assay is used as a transcriptomic profiling method for ranking and prioritizing possible toxicants. Simultaneously, high-content analysis is used as a more specific targeted method, serving as a second-tier toxicity identifier for validation of the omics-based AOP-linked hypotheses. Such methodologies may be able to reduce the need for large-scale data for plenty of NMs required to cover broad dose ranges to ensure high accuracies and apply risk assessment using toxicity pathways.

Nymark *et al.*<sup>192</sup>, proposed a 6-step workflow concerning NM-induced pulmonary fibrosis, resulting in a bioinformatics analysis of several toxicogenomics databases and a data fusion pipeline for AOP-enrichment. This pipeline uses supervised and unsupervised/fixed protocols to both known and novel gene associations (64 genes) and AOP descriptions. Firstly, general schemes describing LF were overviewed, then disease-related information on genes were searched, and interactions between the identified genes were analysed using the GeneMANIA datasets. After that independent genes or genomic groups were analysed for enriched pathways related to functional schemes and they are linked with cellular responses and key events (MIEs, KEs, AEs). The final step was to integrate the generated data into a AOP scheme and a novel, openly accessible AOP-linked molecular pathway representing the effects of MWCTS exposure that uses lung fibrosis in homo sapiens as case study was presented in Wikipathways (<https://www.wikipathways.org/index.php/Pathway:WP3624>). The proposed approach could facilitate connecting the known genomic responses to AOPs and simplify the description of the disease.

In a different study by Nymark *et al.*<sup>193</sup>, toxicogenomics data were considered for analysing and understanding genomic influences of engineered NM exposure. Specifically, the proposed workflows by Kohonen and Grafström (also analysed above) were used as references focusing on cell-based HTS and high-content technology systematically applied in tiered manners for screening potential hazardous effects of engineered NMs. They suggest that such methodologies pose two important advantages: molecular details are provided which support validations of MIEs and KEs and the fact that toxicogenomics data enable the identification of sensitive biomarkers for targeted measurement identified in the AOP. This work associates high aspect ratio materials (HARMs), namely asbestos and MWCNTs, with DNA damage in the form of specific chromosomal fragments. Also, it is suggested that AOPs may implement knowledge on toxic mechanisms for engineered NMs and identify toxicity pathways jeopardising human health. They focus on specific AOPs associated with lung diseases, such as AOP-ID: 173 for lung fibrosis, AOP-ID: 148 for decreased lung function and AOP-ID: 171 for mesotheliomas in rats. Nymark *et al.*, intent to study genomics data in relation to HTS data and disease progression, and eventually create a genomics-driven basis for classifying engineered NMs according to their toxic potencies.

In the work by Halappanavar *et al.*<sup>31</sup>, a vague number of nearly 11000 publications focused on NM toxicity issued between 2000-2013 were screened, and several elimination processes were applied. For example, literature that did not investigate any biological systems were excluded from the study, which ultimately resulted in a selection of 191 studies referring to inflammation and 447 identified records were derived, with no further quality screening. After gathering the desired databases, the different biological events caused by approximately 45 manufactured NMs were identified and analysed in terms of NM type, endpoints, exposure details and post-exposure timepoints. Many endpoints were considered as KEs or even AOs. Tissue injury was the single KE selected for the case study development, since it is plausible, measurable, and regulatory relevant. Moreover, the whole database was re-evaluated in order to search for *in vivo* and *in vitro* measurements. *In vivo* measurements include chronic inflammation, oxidative stress, and histopathology, while *in vitro* measurements include cytotoxicity, oxidative stress, and membrane permeability as biomarkers. It was established that NM-induced toxicity involves an acute inflammatory component, and tissue injury is expected to be used in a case study for AOP development and manufactured NMs risk assessment.

The responses of mice exposed to different NMs, including MWCNTs, nano-TiO<sub>2</sub> and carbon black NMs, were investigated by Halappanavar *et al.*<sup>194</sup>, aiming to create an approach for NM ranking based on their possibility/potency to induce acute lung inflammation. Toxicogenomics data derived *in vivo* were identified with several bioinformatics tools, such as the Agilent Whole Genome microarray platform, and then they were translated into pathway-based BMDs. Transcriptional BMDs were estimated for all expressed genes and pathways associated with lungs exposed to NMs. The potency ranking was made according to the BMD results. Gene and pathway responses were compared between all variants included in the study and a further assessment of their similarities was provided. Even though specific MWCNTs proved to be more potent than other NMs, thus making toxicity testing more urgent for them, NMs did not exhibit vast differences and a clear ranking was not made.

## 5.5 Other models

### 5.5.1 Causal Models

Causal modelling can be used to represent very complex relationships among a set of variables. It is basically a diagram that contains the relationships between dependent and independent variables. A variable's influence on the dependent variable is either direct (direct effect) or transmitted through a different variable (indirect effect). Its ability to handle complexity and portray an issue in a simple way makes it a useful method for NMs' toxicity research and risk estimation.

A few examples of studies using causal modelling are included in this work. Liu *et al.*<sup>106</sup> used an association rule mining approach to identify complex relationships among different types of NP-induced cellular responses. They extracted knowledge from high throughput screening data (HTS), to assist in the discovery of toxicity mechanisms, for the development of bioactivity endpoints and for the establishment of NM structure-activity relationships. Relationships among cellular responses (RAW264.7 and BEAS-2B cells) were investigated through a workflow that included HTS data normalisation and identification of significantly regulated cellular responses. This model was validated through comparisons with existing experimental results. Also, Sizochenko *et al.*<sup>134</sup> used causal inference methods between two nanoQSAR classification models that predict NP toxicity to human BEAS-2B and RAW264.7 cells. The causal interaction of the two nanoQSARs was validated using a random forest model for a selected group of descriptors.

### 5.5.2 Exposure Models

Exposure models are essentially assessment approaches that consider different NM dispersion and exposure routes to humans and the environment. Such models can indicate possible exposure risks and provide help in decision-making and risk-management concerning manufactured NMs. However, there is limited environmental exposure data which complicates model validation for the exposure models or the material flows and makes statistical analysis impossible. Two studies that conducted exposure analyses were assessed in this work.

RedNano is an integrated simulation tool developed by Liu *et al.*<sup>195</sup> for assessing the potential release and environmental distribution of NMs, which considers NMs' mechanistic intermedia transport. Major exposure pathways are identified, and a multimedia compartmental model is developed aiming to estimate environmental NM concentrations and release rates in specific scenarios. Zheng *et al.*<sup>196</sup> developed an environmental exposure model as well, by creating a size-specific, dynamic, probabilistic material flow analysis (ss-DPMFA). This analysis considers particle size, crystalline forms, and coating materials, parameters that potentially influence the material's fate, transport, and toxicity, and conducts a TiO<sub>2</sub>-specific hazard assessment.

### 5.5.3 Dose response

A different methodology was followed by Liu *et al.*<sup>197</sup>, who used dosimetry analysis on different cell lines (EC<sub>50</sub> of human and murine cell lines) to rank MeOx NMs for hazard assessment. Initially, a sedimentation model (SP2N) was developed based on the motion of particles by Brownian diffusion for the calculation of the delivered dose. Then, the obtained *in vitro* hazard ranking was compared to the ranking based on the administered dose, and dose-response analyses were conducted to establish toxicity metrics.

## 5.6 Models and methods implemented as user-friendly tools

The development of precise models that facilitate the screening and filtering of unfavourable candidate NMs is prompted by the increasing demand for evaluating the characteristics and undesirable activity of NMs before spending time and resources on their synthesis and testing. The growing variety of approaches and tools for the computational assessment of NMs negative effects and properties that have been built over the past years serves as a good example of this. However, researchers in the field of nanosafety (such as experimentalists, regulatory specialists, etc.) who could directly benefit from the use of *in silico* methodologies on a daily basis, may be discouraged by the programming environments due to their lack of a solid background in data science, statistics, or programming, or due to time restrictions that may prevent them from learning to develop their own scripts and models. Following this need, several of the generated models -assessed in this study- are already available with a GUI as online applications or stand-alone software. The various technical components of the models are

streamlined using a GUI especially created for non-informatics experts, presenting only the essential components for executing the models. As a result, the group of scientists and professionals who can benefit from the use of computational models in the field of NMs risk assessment or novel NMs design is greatly expanded. The disseminated models will serve as an important tool in future computer-aided NM design and quality control. Table 7 presents a list of models for which a GUI is provided. A more comprehensive presentation of these tools is presented in Appendix II.

Table 7: List of disseminated models and *in silico* methodologies as web applications or stand-alone platforms.

Tool	Availability
A safe-by-design tool for functionalised nanomaterials <sup>70</sup>	<a href="http://www.enalosccloud.novamechanics.com/EnalosWebApps/CNT/">http://www.enalosccloud.novamechanics.com/EnalosWebApps/CNT/</a>
Apellis <sup>34</sup>	<a href="https://apellis.jaqpot.org/">https://apellis.jaqpot.org/</a>
BioDaph	<a href="https://biodaph.cloud.nanosolveit.eu/">https://biodaph.cloud.nanosolveit.eu/</a>
Cytotoxicity (cell viability) prediction for metal oxide NPs <sup>155</sup>	<a href="https://cellviability.cloud.nanosolveit.eu/">https://cellviability.cloud.nanosolveit.eu/</a>
DeepDaph <sup>99</sup>	<a href="https://deepdaph.cloud.nanosolveit.eu/">https://deepdaph.cloud.nanosolveit.eu/</a>
Ecotox models <sup>75</sup>	<a href="https://ecotox.cloud.nanosolveit.eu/">https://ecotox.cloud.nanosolveit.eu/</a>
Enalos QNAR iron oxide toxicity platform <sup>65</sup>	<a href="http://www.enalosccloud.novamechanics.com/EnalosWebApps/QNAR_IronOxide_Toxicity/">http://www.enalosccloud.novamechanics.com/EnalosWebApps/QNAR_IronOxide_Toxicity/</a>
Facet cytotoxicity prediction <sup>88</sup>	<a href="https://facetcytotoxicity.cloud.nanosolveit.eu/">https://facetcytotoxicity.cloud.nanosolveit.eu/</a>
GUIDEnano	<a href="https://tool.guidenano.eu/">https://tool.guidenano.eu/</a>
INSIdEnano <sup>167</sup>	<a href="http://inano.biobyte.de/">http://inano.biobyte.de/</a>
Logistic regression model for the toxicity classification of SPIONs <sup>76</sup>	<a href="https://app.jaqpot.org/model/DcWnWFp9GESI16R4o2av">https://app.jaqpot.org/model/DcWnWFp9GESI16R4o2av</a>
Lung exposure dose calculator <sup>15</sup>	<a href="https://lungexposure.cloud.nanosolveit.eu/">https://lungexposure.cloud.nanosolveit.eu/</a>
Metal oxide nanoparticles cytotoxicity classification	<a href="https://cytotoxicity.cloud.nanosolveit.eu/">https://cytotoxicity.cloud.nanosolveit.eu/</a>
MS <sup>3</sup> bD zeta potential predictive Model <sup>1</sup>	<a href="http://www.enalosccloud.novamechanics.com/nanocommons/mszeta/">http://www.enalosccloud.novamechanics.com/nanocommons/mszeta/</a>
Nano-lazar <sup>38</sup>	<a href="https://nano-lazar.in-silico.ch/predict">https://nano-lazar.in-silico.ch/predict</a>
Nano Protein Corona model <sup>36</sup>	<a href="http://www.enalosccloud.novamechanics.com/nanocommons/NanoProteinCorona/">http://www.enalosccloud.novamechanics.com/nanocommons/NanoProteinCorona/</a>
NanoBio <sup>198</sup>	<a href="https://nanobio.cloud.nanosolveit.eu/">https://nanobio.cloud.nanosolveit.eu/</a>
NanoMixHamster <sup>101</sup>	<a href="https://nanomixhamster.cloud.nanosolveit.eu/">https://nanomixhamster.cloud.nanosolveit.eu/</a>
Nanoinformatics model for zeta potential prediction <sup>89</sup>	<a href="http://www.enalosccloud.novamechanics.com/EnalosWebApps/ZetaPotential/">http://www.enalosccloud.novamechanics.com/EnalosWebApps/ZetaPotential/</a>
NanoInhale <sup>15</sup>	<a href="https://nanoinhale.cloud.nanosolveit.eu/">https://nanoinhale.cloud.nanosolveit.eu/</a>
NanoPot <sup>198</sup>	<a href="https://nanopot.cloud.nanosolveit.eu/">https://nanopot.cloud.nanosolveit.eu/</a>
NanoProfiler <sup>199</sup>	<a href="https://sites.google.com/site/dtclabnp/">https://sites.google.com/site/dtclabnp/</a>
NanoSafer	<a href="http://www.nanosafer.org/">http://www.nanosafer.org/</a>
NanoSerpA <sup>200</sup>	<a href="https://www.cyc-ingenieros.com/nanoserpa/">https://www.cyc-ingenieros.com/nanoserpa/</a>
NanoToxRadar <sup>201</sup>	<a href="https://nanotoxradar.kitox.re.kr/">https://nanotoxradar.kitox.re.kr/</a>
PBPK models and integration with the occupational exposure model <sup>15</sup>	<a href="https://exposurepbpk.cloud.nanosolveit.eu/">https://exposurepbpk.cloud.nanosolveit.eu/</a>
Prediction of MNPs uptake in PaCa2 cancer cells <sup>44</sup>	<a href="http://www.enalosccloud.novamechanics.com/EnalosWebApps/QNAR_PaCa2/">http://www.enalosccloud.novamechanics.com/EnalosWebApps/QNAR_PaCa2/</a>
QSAR for nano-mixtures <sup>130</sup>	<a href="https://kriktcsr.shinyapps.io/TiO2_Dmagna/">https://kriktcsr.shinyapps.io/TiO2_Dmagna/</a>
QsarDB models <sup>202</sup>	<a href="https://qsar.db.org/repository/handle/10967/119">https://qsar.db.org/repository/handle/10967/119</a> and <a href="https://qsar.db.org/repository/handle/10967/214">https://qsar.db.org/repository/handle/10967/214</a>

Quantitative read-across (Read-Across-v4.1) <sup>59</sup>	<a href="https://sites.google.com/jadavpuruniversity.in/dtc-lab-software/home#h.bus7xy9m3cbd">https://sites.google.com/jadavpuruniversity.in/dtc-lab-software/home#h.bus7xy9m3cbd</a>
Random forest model for the genotoxicity classification of MWCNTs <sup>77</sup>	<a href="https://app.jaqpot.org/model/THPwkjY8Oz7yaIFNAYJR">https://app.jaqpot.org/model/THPwkjY8Oz7yaIFNAYJR</a>
RiskGONE <i>in vitro</i> dosimetry tool <sup>203</sup>	<a href="http://www.enalosccloud.novamechanics.com/riskgone/InVitroDosimetry/">http://www.enalosccloud.novamechanics.com/riskgone/InVitroDosimetry/</a>
Sbpot model <sup>198</sup>	<a href="https://sbpot.cloud.nanosolveit.eu/">https://sbpot.cloud.nanosolveit.eu/</a>
Tool for assessment of human exposure to nanomaterials <sup>15</sup>	<a href="https://aerosol.cloud.nanosolveit.eu/">https://aerosol.cloud.nanosolveit.eu/</a>
toxFlow <sup>37</sup>	<a href="https://toxflow.jaqpot.org/">https://toxflow.jaqpot.org/</a>
vythos <sup>35</sup>	<a href="https://vythos.jaqpot.org/">https://vythos.jaqpot.org/</a>
ζ-regression <sup>198</sup>	<a href="https://zetapot.cloud.nanosolveit.eu/">https://zetapot.cloud.nanosolveit.eu/</a>

## 6 Interviews with experts from Academia and Industry

Based on the findings and information gathered from the literature review, we contacted the relevant experts to get their professional opinion about existing grouping and read-across frameworks and *in silico* models. Information was collected on the state-of-the-art of predictive tools and models development (including the models' validation and applicability). Existing tools and models of relevance to the project not identified in the previous tasks have been collected as well. In addition, the experts' opinion on the reliability of the existing *in silico* tools was sought, including the possible data gaps or missing steps to be applicable for research and regulatory purposes. Finally, the potential paths for future improvements in this area were also identified to make the tools regulatory compliant and usable.

The aim of the interviews was to collect information on the following topics:

- The reliability of the grouping approaches, frameworks and models published or publicly available and their relevance for regulatory applications,
- The data gaps or missing steps to be applicable for research and regulatory purposes,
- The development needs to further improve this area and make these regulatory compliant and usable,
- The identification of any missing frameworks, tools and/or models of relevance to the project.

### 6.1 Identification of the experts

The research group initiated the identification of experts and stakeholders from academia, industry, and regulatory authorities for the surveys as early as possible. To create a first pool of experts a list of key authors (e.g., first and last author/corresponding author) from the publications collected during the review was composed. Besides the experts identified through the literature search, we also used our extended network of commercial work partners by participating at the EU NanoSafety Cluster (<https://www.nanosafetycluster.eu/>), the Nanotechnology Industries Association (NIA, <https://nanotechia.org/>), and BioNanoNet (<https://www.bnn.at/>). Furthermore, from our participation in various H2020/HORIZON Europe projects including on NMs fate, risk governance, development of tailored nanoinformatics tools, SbD, and Safe Innovation Approaches, grouping and read-across, it was possible to contact the nanoinformatics experts from the different consortia.

Experts were contacted in different ways:

- By e-mail to the key authors identified from the collected scientific publications,
- By personal contact of the network collaborators,



- By posting an open call/invitation to LinkedIn, and
- By the BioNanoNet and NIA newsletter to their list of partners/collaborators.

## 6.2 Organisation, preparation, distribution of the questionnaires and collection of the data

The procedure devised for distributing the questionnaires was designed to maximize the survey's impact and ensure that any filing weariness would be prevented. This procedure comprised:

- Set clear, attainable survey goals identified from the data gathered during the review,
- Define a clear set of questions per topic to cover the required outcomes,
- Use the proposed questions to draft more analytical questions that were customised and on the purpose of data,
- Pick the best questions that would lead to data retrieval maximisation at the optimum time frame (max 30 minutes for online questionnaires, and preferably no more than 15 minutes),
- Craft questions to maximise the acquired results, with an optimal ratio of closed (e.g., yes/no, multiple choice, agreement/disagreement - Figure 28) to open-ended (e.g., free-text, critical, opinion - Figure 29) questions,
- Ordering the questions so that the most significant questions are answered first, as there was a chance that people started, but did not finish the questionnaire ensuring all answers are captured.

Great care was taken to ensure that the questions were as short and as simply framed as possible. To make sure that no prejudice entered the questions, they were evaluated numerous times. An example of biased and unbiased questions under the same context is:

- Biased: Do you think that the currently available data inhibit the development of validated models for use in the effective and reliable testing of NMs?
- Unbiased: What are the barriers, if any, in your opinion, for the development of effective and reliable models for the testing of NMs?

Are you models freely available or is there a fee for using?

- ☐ Free  
☐ Freeware  
☐ One-time fee  
☐ Subscription

Are you providing the source code of the tools you are developing?

- ☐ Yes  
☐ No

Do you provide APIs for remote accessing of the *in silico* tools?

- ☐ Yes  
☐ No

Are you aware of any *in silico* methodologies integrated in industrial, regulatory or research level for the SbD, grouping, and read-across of nanomaterials?

- ☐ Yes  
☐ No

Figure 28: Example of multiple-choice questions.



What are the difficulties in the integration of the results of the *in silico* methodologies in the SbD of novel nanomaterials or regulatory applications?

How do you think it would be possible to reinforce the confidence of the stakeholders in reliability of modelled results and in *in silico* methodologies?

Which are the key parameters that future *in silico* investigation of nanomaterials should focus (on)?

Figure 29: Example of free-text, opinion questions.

In all cases, participants were clearly informed regarding the scope and desired outcomes from the questionnaire, ensuring anonymity of responses and the handling of confidential information. These included:

- All experts were given the opportunity to fill in the questionnaire online, using a secure environment, fill in a Word version of the questionnaire offline or answer the questions via a live remote interview.
- All experts were informed that the questionnaires would be shared with ECHA in confidentiality and all published responses would be anonymised removing any personal or commercial data.
- There was the provision for questions to be flagged as confidential. In this case, the responses and data are separated from the non-confidential data and shared with ECHA in the annex of this report but will not be mentioned or published. They are used for analysis and reaching conclusions.

The questionnaires were put together in close collaboration with ECHA, based on the data collected and critical analysis from the review. Questionnaire duration did not exceed 30 minutes, with an optimum duration of around 15 minutes. To retain the participant's attention on the questions, the questionnaire was divided into five sections, with each one focusing on a different topic:

- Section 1 - Personal information,
- Section 2 - *In silico*, grouping, and read-across methodologies,
- Section 3 - Data availability,
- Section 4 - Dissemination,
- Section 5 - Gaps and future steps.

### 6.3 Risk mitigation measures

The inclusion of external stakeholders presents inherent uncertainty for conducting surveys/interviews and collecting the relevant information because timely and effective collaboration is challenging to guarantee. The time schedule was organized effectively, making use of project management expertise and resources to ensure prompt implementation and achievement of the desired outcomes. The management team took precautions when structuring the study to mitigate against risks related to limited stakeholder participation to minimize the impacts and assure low chance of risk manifestations during the implementation of the tender (

Table 8).

Table 8: Identified risks and risk mitigation measures for the surveys.

Identified risks	Severity / Possibility	Mitigation measures
<b>Low participation hinders the richness of the obtained data</b>	High/Medium	<i>Preventive:</i> Establish contacts immediately upon contract award, enrich experts' database, establish good practices for communication, raise trust and experts' engagement. <i>Corrective:</i> Use existing networks to identify alternative contact persons. Offer easy formats for collaboration.
<b>Stakeholder groups are unbalanced (e.g., absence of representative samples of a group, over-representation of samples of a group)</b>	Medium/Low	<i>Preventive:</i> Equal number of members of expert groups will be contacted and invited for participation. <i>Corrective:</i> The organisation of the experts' database will allow early identification of under/over-representation so that expert groups can be balanced by inviting extra members or by adjusting the weighting of larger groups.
<b>Delays in receiving replies</b>	High/Medium	<i>Preventive-Corrective:</i> Regular reminder e-mails will be sent to the experts' network to ensure that they participated to the survey.
<b>Length of survey/Incomplete surveys submission</b>	High/Medium	<i>Preventive:</i> Engagement of experts can be heavily influenced by the time needed to complete the surveys. The design of the questionnaire will ensure that the survey will not require more than 30 min to be completed. <i>Corrective:</i> Information will be confirmed with experts during the initial tasks and the survey length can be reduced if deemed necessary for maximising participation.

The full questionnaire can be found in the Appendix III of this report. The questionnaire was uploaded to EUSurvey (<https://ec.europa.eu/eusurvey/home/welcome>) an online survey management tool for producing and making public forms available. The survey can be found via this link: <https://ec.europa.eu/eusurvey/runner/InSilicoReadAcrossApproaches>

The online survey regarding the modelling tools and read-across approaches for NMs hazard and risk assessment reached approximately 150 experts. The results collected from 11/11/2022 to 19/01/2023. The experts' answers were analysed and synthesised under a single report, providing quantitative data whenever possible. They are presented in the next sections, following the same structure used in the distributed questionnaire. The responses of closed questions are also presented in graphical format in the Appendix IV of this report.

#### 6.4 Experts' information

At the reporting period, 36 experts responded to the survey, the majority of which (61%) participate in EU, national, or international research (nano)material related projects (Figure 107). In addition, 56% of the contacted experts are developing or involved in the development of models and *in silico* methodologies for NMs (Figure 111). As far as the sector they come from is concerned, the greatest part of the respondents (69%) is involved in Academia, and the rest of them are occupied in the Industry/Enterprise sector (Figure 108).

## 6.5 *In silico*, grouping, and read-across methodologies

According to the submitted responses, most of the experts (64%, Figure 109) are aware of existing grouping and read-across frameworks such as the Read-Across Assessment Framework (RAAF) supported by ECHA, the general framework outlined in the OECD documents, the GRACIOUS framework and the DF4nanogrouping framework. Many of the experts reported tools and methodologies as grouping/read-across frameworks (e.g., the OECD QSAR toolbox, the AOPs, the US EPA, the Jaqpot, the Enalos, the Apellis, the ToxTree tools, the NanoSolveIT and the SbD4Nano projects). A definition was also provided “the framework for grouping and read-across assumes that similar structures have similar activities. Based on this and an appropriate justification, data gaps for a target chemical can be filled with already existing information of its analogues”.

69% of the responders claimed to have employed *in silico* methods or tools for research or regulatory purposes (Figure 110) as follows:

- For research:
  - In different QSAR modelling studies (development of new methods for validation of regression and classification models and development of variable selection algorithms, and numerous applied and comparative studies),
  - In QSAR studies for the prediction of toxicity of NPs against various endpoints,
  - For simulations of materials, processes, and complex systems (e.g., development of multiscale modelling methods based on the principles of statistical mechanics to predict the dispersion of NPs in polymers as well as rheological, mechanical, and permeability properties of the resulting nanocomposites),
  - For the prediction of sorption and diffusion in nanoporous materials, such as zeolites,
  - In the risk assessment of NMs as well as of solvents (e.g., deep eutectic solvents),
  - For data gap-filling of toxicity endpoints required under chemical regulations and for regulatory compliance for K-REACH (The Act on the Registration and Evaluation of Chemicals of South Korea),
  - For the development of toxic-genomics data analysis tools (e.g., NEXTCAST, ECOSAR),
  - For relevant data analysis tasks such as data collection and curation,
  - For molecular docking applications,
  - When drafting the QSAR Report Formats (QPRF) and Read-Across based expert opinions on chemical's properties/activities,
  - In different research projects.
- For regulatory purposes:
  - For grouping of toxic chemicals,
  - In the chemical registration and evaluation in national level (e.g., in Korea),
  - Advising regulatory agencies on (chemical) hazards (e.g., Australian agency NICNAS).
- For commercial purposes:
  - For the development of *in silico* methods in companies,
  - In different collaborative research projects,
  - For the development and application of computational models for regulatory purposes (in different regulatory frameworks such as REACH and ICH).

The reported tools in this course are the OECD QSAR toolbox, the Discovery studio, R programming language, and the Chemistry Development Kit. The majority (56%) of the already developed *in silico* methodologies are reported as dataset dependent (Figure 112). The experts were also asked to provide information on the models and *in silico* methodologies they have developed or have been involved with developing:

- Development of specific QSAR models for NMs with a focus on REACH-relevant properties.

- Development of QSAR models for nanoforms to toxicologically relevant endpoints such as cytotoxicity, activity against bacteria and *Daphnia*, based on machine learning (ML) algorithms.
- Development of multitasking QSAR classification models (mtk-QSAR).
- Employment of multi-objective optimisation techniques including Monte Carlo optimisation and artificial bee colony algorithms.
- Development of polymer self-consistent field-based methods in one and three dimensions for the prediction of the free energy of NP-polymer systems in dependence of the spatial arrangement of NPs for the purpose of optimising NP dispersion in the matrix.
- Development of strategies based on systematic coarse-graining, along with connectivity-altering Monte Carlo and molecular dynamics simulations at atomistic and coarse-grained levels, for the prediction of glass transition temperatures, elastic constants, and permeabilities of polymer-matrix nanocomposites.
- Development of Brownian Dynamics/kinetic Monte Carlo simulation methods for the prediction of linear and nonlinear rheological properties of entangled long-chain polymer melts and dispersions of NPs therein.
- Development of toxic-genomics based models in the framework of AOPs and development of quantitative AOPs using Bayesian model based on experimental nanotoxicity data.
- Mixture toxicity prediction for chemical mixtures and nano-mixtures.
- Computational modelling of physicochemical properties and descriptors tools (e.g., atomistic nanodescriptors).
- Development of multiscale models for adsorption of biomolecules on NMs surfaces (bio-nano interactions).
- Development of user-friendly tools for the *in silico* assessment of NMs (e.g., Jaqpot, NanoToxRadar, NanoProfiler, NEXTCAST) and of decision support tools for the safety assessment and regulatory compliance of NMs. The reported tools available through a GUI are presented in Appendix II.
- Development of *in silico* models as part of scientific projects (e.g., SbD4nano, NanoSolveIT, NanoBridges).

The experts involved in modelling, are developing classical QSAR, nanoQSAR, nano-mixture QSAR and multi-tasking QSAR models, grouping, read-across and read-across structure–activity relationships (RASAR) models, PBPK models, mechanistic models, mixture toxicity prediction models and (quantitative) AOPs. Different modelling techniques were reported, such as random forests, neural networks, Bayesian models. Also, five out of the total 36 experts are involved in physics-based modelling (e.g., *ab initio* simulations) often coupled with ML.

The reported standard validation methodologies that are employed for the developed *in silico* methodologies are in many cases in line to the recommended OECD guidelines for validating QSAR models and include internal and external validation, and the Y-scrambling test. The internal validation is employed especially in cases of limited data and consists mainly of a k-fold cross-validation scheme. Specifically, four experts report using k-fold (primarily 10-fold) cross validation. In external validation, datasets are split into training and test sets (in some cases, in one additional external set), where the training set is used for model development and selection, while the test/external sets are exclusively used for model validation. In the case of physics-based models, comparison against experiments in well-defined systems is performed.

Furthermore, several deterministic and probability-based metrics are employed to perform validation. In case of regression the following metrics are reported:

- Coefficient of determination ( $R^2$  and adjusted  $R^2$ ),
- Mean absolute error (or deviation) (MAE),

- Mean square error (or deviation) (MSE),
- Negative mean absolute error (NMAE),
- Spearman's correlation coefficient,
- Standard error (SE) and standard error of estimate (SEE),
- Standard deviations of errors of prediction (SDEP) and root mean square error of prediction (RMESP),
- Other metrics (not defined):  $Q_{LOO}^2$  or  $Q_{LMO}^2$ ,  $Rm^2$ ,  $Q_{F1}^2$ ,  $Q_{F2}^2$ ,  $Q_{F3}^2$ .

One of the experts also proposed the acceptability criteria presented in Table 9. We must note at this point that a universal rule of thumb for the appropriateness of  $R^2$  values does not exist. The type of the variables included in the model, how the variables are measured, and how the data is transformed are just a few of the factors that affect how accurate a statistical measure will be. The  $R^2$  itself neither provides any information on the causal relationship between the descriptors and the endpoint nor indicates the level of “correctness” of the model. As mentioned in the §5, a combination of statistical measures should be examined, as well as other tests should be applied to assess the data quality and the robustness of the model. Finally, the interpretation of the model is equally important.

Table 9: Acceptability criteria based on  $R_{CV}^2$ .

Criterion	Result
$R_{CV}^2 > 0.95$	Excellent
$0.85 < R_{CV}^2 < 0.95$	Very good
$0.80 < R_{CV}^2 < 0.85$	Good
$0.75 < R_{CV}^2 < 0.80$	Satisfactory
$0.70 < R_{CV}^2 < 0.75$	Acceptable
$R_{CV}^2 < 0.70$	Not acceptable

In case of classification the following metrics are used (mainly derived from the confusion matrix):

- Accuracy,
- Precision,
- Sensitivity,
- Specificity,
- Selectivity,
- F-measure,
- MCC,
- Balanced accuracy,
- Receiver operating characteristic (ROC) curve - area under the curve (AUC),
- Wilks's lambda ( $\lambda$ ),
- Chi-square ( $\chi^2$ ),
- p-level.

The reliability of the predictions is ensured in various ways. The use of curated datasets, and of large and balanced validation sets is the basis for the development of a reliable model. Next, the favourable validation metrics (e.g., metrics over 0.7) are indications of credible models. In addition to the statistical metrics used for the validation of the models, their applicability domain definition is equally important, as stated by six experts. Virtual predictions of e.g., ecotoxicity and cytotoxicity of new materials, comparisons to existent data from the collaborators' network or with literature data and in general post-hoc experimental testing of predictions (e.g., measurement, cross-lab differences, biases, model

systems, comparisons with the statistical modelling error, uncertainty estimation, etc.) is contributing to the reliability of the developed models. Another way to ensure the reliability of the predictions is to apply intelligent consensus prediction in case that multiple good models for the same endpoint exist<sup>204</sup>. Finally, in the case of physics-based models, the reliability is ensured by comparing against experimental measurements from a variety of systems. In detail, when developing “bottom-up” physics-based modelling methodologies, predictions from different levels of representation (atomistic, coarse-grained, mesoscopic) for a given system are evaluated to agree, with a minimum of adjustable parameters.

To assess the applicability domain of the developed methods, the plurality of experts use -depending on the type of endpoint (regression/classification) and the type of model (linear/non-linear)- the leverage approach, the Euclidean distance method, the standardisation approach, the confidence estimation approach, the *k*NN approach, the range-based approach, and the multi-dimensional evaluation based on the knowledge graph or they use custom made approaches coupled with uncertainty estimation. The opinion was expressed that in addition to the numerical calculation of the applicability domain, the nanoforms require a specific assessment based on the range of NMs used in the database. This includes the requirement of specific information (such as size, coating, etc.) and the limitation to chemical families (such as metal-oxides, quantum dots, metals, etc.). Different methods of feature selection and successive principal components analysis are also used for directly analysing properties of the multivariate descriptor space of the training and validation compounds. The testing on diverse datasets, on realistic scenarios related to use-case domains, and the application to a variety of systems, often in collaboration with industry is also performed to assess the applicability domain of the developed methodologies. Another opinion expressed that in case of small datasets the applicability domain is not routinely assessed.

To interpret the developed model, nine experts reported that they study the selected descriptors and their contribution to the predictions. In linear models this is performed by analysing the standardised coefficients as measures of the relative importance-contribution of the descriptors in the model. In non-linear models, this is more difficult because feature importance is local rather than global. In such cases the descriptor analysis is performed using the sensitivities and mean values of the different descriptors in the model (e.g., remove one descriptor at a time, and check performance variation). In both cases, it is very important that the selected descriptors participating in the model are easy to understand and explain/interpret. In this course, the literature review can provide some evidence for understanding or assuming potential mechanisms of toxic action, as well as the comparison with other methods. Furthermore, simulation results are extensively post-processed to reveal mechanistic information about the phenomena and properties of interest. Finally in cases of AOPs and mode-of-action models the mechanistic interpretation is derived by the model itself.

Experts involved in model development are not always providing model documentation using a specific template along with their models (45% of the experts are providing a standardised report, Figure 113). In case they do, they usually include a QMRF (QSAR model reporting format) report and, less frequently, a MODA (modelling data) report, a QPRF (QSAR prediction reporting format) or a custom-made report.

## 6.6 Data availability

Most experts' opinion (69%) on data availability was that barriers do exist (Figure 114). The 68% of these experts emphasised three data availability aspects: quantity, sparsity, and accessibility. Standardised, structured, annotated, and reliable nano-related data are limited in comparison to non-nano chemicals, and they are not always accessible (e.g., data obtained in industrial level is often confidential and not available in the public literature). Furthermore, there is no central repository to

host the data like in the case of chemical data. Data are found in different formats and there are consistency deficiencies between the data presented in different sources, including which parameters are considered relevant to identify a nanoform and how these are measured/calculated (e.g., duration of exposure, aggregation rate). The definition of NMs is sometimes incomprehensible (e.g., it was stated that in some cases NMs are confused with microplastics). Data are also difficult to retrieve in bulk format (e.g., using application programming interfaces, APIs<sup>c</sup>) to build automated workflows. Furthermore, it was highlighted by at least two experts that in some cases, data for the same endpoint are collected in different experimental conditions. The resulting data heterogeneity is due to the lack of an established test guideline for NMs, or to the absence of universally agreed ontologies, even from the European Commission who was unsuccessful in creating a meaningful Open Data Pilot (e.g., lack of clear definition of what is and is not data, data is very easy to be excluded from the Open Data expectations, default agreements are not compatible with Open Data practices). Also, it was expressed that stakeholders do not understand how and why data can or will be reused and some are afraid of data misinterpretation.

Experts' opinion was divided on the availability of datasets of sufficient quality for developing *in silico* methodologies (Figure 115). The experts that agreed on the availability of such datasets (47%), claimed in their majority (59%) that they are also easy to access (Figure 116). The bottlenecks regarding the access of such datasets are due to the lack of systematically organized, curated, and homogenous open data sets (e.g., deposited in repositories). Paywalls may apply and private databases still exist. To address the abovementioned barriers initiatives are in place to create repositories with ready-to-modelling data such as the NanoPharos database.

To build their models, experts in most cases (15 out of 19 experts) use published literature datasets or data found in public databases such as the eNanoMapper ("Nanosafety Data Interface"), the NanoCommons database, the nanoPharos, the PubChem, the ChEMBL, the BindingDB and the EPA (U.S. Environmental Protection Agency) databases, and the toxicity databases available in QSAR ToolBox. In other cases, experts use the data shared from experimental academic research groups whose work they trust or their consortia collaborators. In one case, the use of custom datasets obtained by either experimental or computational methods was reported.

Data quality control is a standard procedure prior to any modelling activities for the majority of the experts (70%, Figure 117). Datasets are curated/cleaned by manual and automatic procedures, including the elimination of duplicates, of incomplete data, and of data with inconsistent results. In addition, experts perform source vetting, meta-data tracking, compare the data on the same system from different groups to the extent possible and use chemical and biological data curation according to the publication of Fourches *et al.*<sup>205</sup>. Other experts reported that they check for unit consistency, very sparse columns or columns that include fixed values, and calculate correlations between input variables. Data standardisation was also reported as a data quality control procedure. Finally statistical tests are reported to be performed for data quality control such as the tau test for outliers, test of homogeneity, and other omits-based metrics.

Experts prefer both FAIR (findable, accessible, interoperable, and reusable) and Open Data<sup>d</sup> to develop their models, whereas some of them do not have any preference (Figure 118). Regarding the type of properties used for modelling purposes, the physicochemical properties are highly preferred, followed by (eco)toxicity and biological data (*in vivo*, *in vitro*). Other properties include atomistic, molecular, and

---

<sup>c</sup> An API is a set of standardised rules that enable various applications to communicate with one another, as an intermediary layer, which handles data transfers across systems.

<sup>d</sup> Open data is available without restriction, whereas FAIR data is accessible under well-defined conditions.



periodic table-based descriptors, as well as pathway and chemical protein interaction data, biocidal and drug-related properties. Nonetheless, six of the experts stated that there is no preference in a particular data type, as all data is essential.

The responders agreed that data collation from different sources should be performed in case of data scarcity (Figure 119). Nonetheless, in this data collation process barriers should be taken into account such as the data compatibility and consistency when dealing with different data sources. In fact, the insufficient meta-data regarding the experimental conditions, methods, and protocols employed during the activity/toxicity testing, the differences in data quality procedures, and the lack of standardised test and procedures to obtain curated collections of data of adequate size create ambiguities and thus, it is difficult to ensure that the data from different data sources refer to precisely the same systems. In addition, due to the multicomponent description of the nanoforms, it is not easy to recall lists of compounds and to create a collection of compatible NMs to develop a model. In addition to data heterogeneity, data gaps and unbalanced data in terms of quantity and quality create barriers for data collation. The opinion was also expressed that existing modelling methods can't cope with data from different sources and thus, new methods need to be devised. Nonetheless, suggestions on the methods were not provided. The differences in the copyright and license terms of the different sources should be also considered before data collation.

## 6.7 Dissemination

Regarding the dissemination of the developed *in silico* methods as user-friendly workflows, most of the experts (78%) agree that it is necessary for the support of regulatory applications and the SbD of novel NMs (Figure 120). In this course, the ones who develop or already developed such methodologies - depending on the orientation of the software- provide them as web applications or tools, stand-alone/locally installed software or they provide the workflows in public repositories (GitHub, university repositories or other libraries) linked to the relevant publication. In addition, the source code of these tools is commonly provided (55% of the experts, Figure 123), however, this is less common for the corresponding APIs (40% of the experts, Figure 124). Finally, the experts involved in the development of computational methods usually provide the tools as free or freeware<sup>e</sup> software and in five cases under a subscription scheme (Figure 122). In cases of a subscription a use license is also always provided nonetheless, a licensing system is not always provided for free tools (Figure 121).

The experts were also invited to comment on the level of integration of the developed *in silico* methods. Most of the responders (56%) claimed that they are not aware of any *in silico* methodologies integrated in industrial, regulatory or research level for the SbD, grouping, and read-across of NMs (Figure 125). Among the known already integrated *in silico* methodologies QSAR (such as classification methodologies, artificial neural networks – ANNs), quantitative adverse outcome pathways (AOPs) and the GUIDEnano tool were reported. More specifically, in research level the NanoSolveIT, NanoinformaTIX, SbD4Nano projects dealing with the NMs SbD, grouping, and read-across. It was also mentioned that the Australian regulatory agency NICNAS uses ML-based models. Finally, QSAR and read-cross methodologies for chemicals are acceptable by REACH, however this is not the case for the NMs (guidelines for their usage in nanoforms are not completed and published yet).

The confidence of relevant stakeholders regarding the reliability of the developed models and *in silico* methodologies was characterised in general as quite low-medium or insufficient for regulation

---

<sup>e</sup> Free software allows its free-of-charge use, but also gives the users the right to study, modify/change and distribute the original or modified versions of the software. To achieve this, access to the software's source code is considered a prerequisite. On the other hand, freeware software can be also used for free, but it cannot be modified or redistributed.

compliance. It was also reported that the stakeholders' confidence depends on the application and that researchers tend to have more confidence than industrial/regulatory stakeholders. Some of the experts also claimed that they are not aware of the level of confidence of the stakeholders to the developed tools, due to lack of feedback from their part. Finally, the opinion was expressed that although there is a potential to develop robust tools, the lack of practices (including compliance with FAIR principles for data and models or community standards for how to share them, transparency on the physical principles and approximations invoked) prevents stakeholders from using them on a routine basis.

## 6.8 Gaps and future steps

Considering the gaps on the NMs *in silico* investigation, the main issue that eleven experts underlined is the lack of sufficiently good, complete, FAIR, and representative datasets for the development and validation of reliable models and methods. Examples of data-related gaps communicated by the responders are, the scarcity of wide databases with a significant amount of standardised endpoint values for a range of different nanoforms and the lack of reliable and accurate experimental data (and meta-data) from well-defined systems (e.g., the lack of standardised methods to characterise NMs, physicochemically and biologically may possibly result in extremely noisy data). The lack of specific, efficient, and fast-to-calculate descriptors to obtain a numerical description of inorganic NMs with easier mechanistic interpretation, the unstandardised ontology, and the lack of data harmonisation were also reported by three experts.

Apart from the data-related issues, experts exposed the problems of confidence of the stakeholders on the developed computational methods, leading to the gap for reliability of *in silico* results between *in vitro/vivo* and *in silico* scientists. This is due to the limited information from the developers (it is mentioned that they need to provide it properly -especially the physicochemical information), the lack of sufficient data to provide a good applicability domain, the absence of systematic vetting of models and their independent judgement, and the need of understanding of the *in silico* limitations for their efficient use. The lack of a strong link of the computational methods with realistic use cases is also reported as a confidence-related gap (e.g., the transfer to the times, costs, quality that manage the industry).

Finally, regarding the development of methods for the NMs *in silico* assessment, some of the responders commented on the development of innovative ML approaches capturing more of the properties of "biologically relevant entity" in the model, e.g., understanding the corona structure and effect on biological responses, could lead to much better prediction of *in vivo* responses. In addition, they commented on the development of reliable multiscale methodologies based on physical understanding and judicious approximations that can take advantage of available computational hardware and ML tools. The groups' definition, the feature importance for non-linear methods and the size of NMs *per se* were also reported as major limits for the development of computational tools.

For the applicability of *in silico* methodologies used for bulk chemicals in the case of NMs, the experts in their majority (56%) believed that they are applicable or are applicable with modifications (Figure 127). The models developed for bulk chemicals will be not applicable for NMs but the basics of QSAR (including multi-tasking QSAR), read-across, and PBPKs for bulk materials are applicable to NMs considering suitable descriptors and meta-data. ML and advanced ML methodologies as well as transport simulations and physics-based methods are equally applicable to bulk materials and NMs.

Most of the proposed modifications (7 out of 11 responses) included the development of suitable nanodescriptors (obtained experimentally or theoretically). It is necessary to consider that nanoforms are not only defined by their formula but require a multicomponent identification. Then it is necessary to use different and specific descriptors that consider the complexity of nanoforms, encoding structural

information (such as size, shape, and aspect ratio), chemical composition of their components (which can be inorganic species, biomolecules, and organic chemicals) and how those components interact (corona formulations and agglomerations). In this course, dynamics according to NMs size may also need to be taken into account. The incorporation of other key aspects such as the experimental conditions should be also considered to study the property of interest thus, multi-tasking QSAR methods might be more useful than classical QSAR. Finally, as mentioned before, the feature importance for non-linear methods should be also examined.

In their majority (72%) the experts agreed that the integration of the results of the *in silico* methodologies in the SbD of novel NMs or regulatory applications is possible (Figure 126). However, they emphasised some of the difficulties in this process -many of them already pointed out in the previous sections. To begin with, the small NM datasets -compared to the datasets of bulk chemicals- lead to the development of models for a limited number of NMs. In addition, it was mentioned that the information required to identify a nanoform should be clarified and which part of this should be reflected in the model nonetheless, regulatory agencies have not yet published the regulatory specifics of the models. As a result, over simplistic models that do not capture all relevant properties of NMs are developed. Methodology trust issues also arise due to the uncertainties (e.g., proper validation, DoA, in QSAR modelling predictions of the target activity are based on the structure which is different between chemicals and NMs) and lack of understanding of the methods from the stakeholders. The lack of different models' web-implementation and the lack of API access are also impeding their integration in the SbD of NMs. Finally, validation of the models results in realistic and extended use cases is missing and *in silico* or *in vitro* correlation with *in vivo* results is considered currently unreliable.

To reinforce the confidence of the stakeholders in reliability of modelled results and *in silico* methodologies, experts mentioned that models should be based on sufficient and qualified datasets, and they proposed to demonstrate their reliability by the application of examples of success in wider libraries of FAIR and Open data, that demonstrate strengths-limitations of each model. Appropriate understanding of models' limitations should be shared with the stakeholders. The produced models must be systematically and thoroughly validated and all information about the model development, validation, etc. should be clearly presented. Moreover, providing a rigorous applicability domain would help the stakeholders to decide if they should trust the model or not. Providing different results estimated from various models would also be useful for the users to assess if there are any consensus results. Honest talk about the limitations of experimental work should be also performed and explicit guidelines from institutions will contribute to standardise the field and improve the confidence of the users. Five of the experts also agreed that a way to reinforce the confidence on the nanoinformatics models is to perform post-hoc experimental testing of the properties of interest to confirm/validate the predictions generated by models and to prove that the *in silico* predictions are accurate and useful. This includes, for instance, the communication of the advantages of using reliable *in silico* methodologies for at least virtual screening purposes and the presentation of success stories involving the design/optimisation of NMs for specific applications highlighting the leap in performance. Finally, the opinion was expressed that to reinforce the confidence of the stakeholders in the computational methodologies, it is necessary to design a transnational research project, and explore the possibilities.

The key parameters that future *in silico* investigation of NMs should focus according to the experts is, to begin with, the collection of good and sufficiently large data sets, capturing all relevant properties of NMs and their distributions (e.g., the NMs lattice structure, physicochemical properties, biological activity and chemical reactivity, incorporation of the experimental conditions such along with their reliability). In addition, new descriptors should be developed that are fast and easy to calculate which can address larger size NPs having more than 1000 atoms. All data should be also curated, checked,

and approved by experimentalists, and well organised as ready-to-model datasets. The transparency should be also ensured and the compliance with the Open Science principles (FAIR and Open data) as well. The NMs systematic encoding should be also supported, by extending, for instance, the InChI and InChIKey representation to the nano level. Effort should be also placed on the development and extensive validation of the obtained models. Models should be developed considering several components and the inherent complexity of NMs. For example, protein coronisation is missing in the model development even though it is significantly impacting on NMs' surface character. Computational efficiency and models' reliability need to be improved so that *in vivo* predictions are enhanced without requiring much more *in vivo* testing. Finally, a link with NMs applications should be established. For example, a computational study on NMs should indicate clear paths for transferring the knowledge gained towards the specific application in which these NMs can be exploited.

## 7 Case studies

The objective of this task was to perform three case studies of grouping, read-across, and *in silico* models applied to NMs based on the findings of the literature review and the results of the surveys/interviews with experts.

The case studies covered a broad range of NMs e.g., Au, Ag, ZnO, MWCNTs, organic NMs, and some novel and advanced NMs like 2D (e.g., graphene), high aspect ratio and multicomponent NMs. The three case studies relate to different aspects of regulatory relevance such as human, or environmental hazard and/or exposure and applications and drug delivery, consumer products, environmentally benign NMs. The case studies tested the applicability of grouping and read-across frameworks and *in silico* models to exposure, hazard and risk assessment of NMs, either with distinct or a combination of chemistries. *In silico* workflows that combine different models were created that can assist with exposure, risk, and hazard assessment, and prediction of NMs properties useful for SbD paradigms.

In the first case study read-across modelling is performed for assessment of the anti-infective activity of various carbon-based nanoparticles (CNPs). Three models were developed using three read-across frameworks and a dataset of interactions between CNPs (fullerenes, CNTs and graphene sheets (GS)) and a SARS-CoV-2 RNA fragment from the literature review.

The second case study summarises the use of different web-applications and tools to assess diverse endpoints for titanium dioxide (TiO<sub>2</sub>) NMs. Here, seven web-tools were employed to predict TiO<sub>2</sub> NM toxicity-related endpoints, physicochemical properties, biodistribution in different ecosystems, and biodistribution in humans after inhalation. The tools used in this case study were obtained from the literature review, and from interviews with the experts.

The third case study was oriented towards the implementation of a Structure–Activity Prediction Network (SAPNet), by combining two nanoinformatics models available in the literature for the prediction of the NMs zeta potential in water and next in potassium chloride (KCl) solutions. The SAPNet methodology, as well as the two combined models, were retrieved during the literature review.

### 7.1 Case study 1: Development of read-across models for the assessment of the anti-microbial activity of various carbon-based nanoparticles

#### 7.1.1 Scope of the first case study

This case study assesses three read-across methods or tools reported in the literature:

- The EnaloskNN method in the KNIME (Konstanz Information Miner) platform as a node with enriched read-across functionalities<sup>1,70,75,88,89,155</sup>.

- The Apellis web application developed for training grouping/read-across models that predict engineered NMs toxicity-related endpoints (<https://apellis.jaqpot.org/>)<sup>34</sup>.
- The Quantitative read-across (Read-Across-v4.1) software that predicts NM biological activity using several similarity based functions (<https://sites.google.com/jadavpuruniversity.in/dtc-lab-software/home>)<sup>59</sup>.

The methods ranked according to their ability to generate reliable models that predict nano-related endpoints, and their ease of use. The various tools were evaluated using the same data so that comparisons between them may be drawn. Comments are also made on their lists of features.

The rationale for the study of the effect of CNPs on a SARS-CoV-2 RNA (in brief cov-RNA) fragment<sup>f</sup> was their high antimicrobial activity and their possible utility in the COVID-19 pandemic. Specifically, the potential of the complexation of the CNPs and the cov-RNA fragment was examined, which might lead to its stabilisation/inactivation. This property may also be of use in antimicrobial applications such as protective coatings. Molecular interactions between CNPs and the cov-RNA fragment have been investigated by Zhang *et al.*<sup>127</sup> using molecular dynamics simulations, where each complex was randomly generated and optimised. They calculated the total potential energies, the Van der Waals energies, and the electrostatic energies of the complex, the isolated CNPs, and the isolated cov-RNA fragment to elucidate the mechanisms of interactions of CNPs with the cov-RNA fragment. The energy of the interaction ( $E_{\text{int}}$ ) provides the driving force for complexation<sup>206</sup>. The energies indicated that CNPs form stable complexes with the RNA fragment (the cov-RNA fragment is adsorbed onto the CNPs).

#### 7.1.1.1 Dataset

The dataset from Zhang *et al.*<sup>127</sup> consisted of physicochemical and molecular descriptors for 17 different types of CNPs including fullerenes, carbon nanoballs, carbon nano-onion, single-walled and MWCNTs, nano-rope, and CNT complexes, monolayer and bilayer graphene. The target variable for the predictive model is the total interaction energy ( $E_{\text{int}}$ ) from molecular dynamics simulations of the CNP with the cov-RNA fragment. Apart from those derived from the simulations, other physicochemical descriptors were calculated such as the molecular weight (MW), the overall surface area (OSA), the volume (V), the specific surface area (SSA), and the sum of degrees of every carbon atom (SDeg). The carbon structural family each CNP belongs to can be also included as a nanodescriptor. Those independent variables along with the calculated target values are displayed in Table 10.

---

<sup>f</sup>A model molecule of frameshift stimulation element from the SARS-CoV-2 RNA genome.

Table 10: Physicochemical descriptors and the total potential energy interaction (endpoint) between the CNPs and the cov-RNA fragment. CNPs with an asterisk belong to the test set.

CNP	Structure	MW [g/mol]	OSA [nm <sup>2</sup> ]	V [nm <sup>3</sup> ]	SSA [m <sup>2</sup> /g]	SDeg	E <sub>int</sub> [kJ/mol]
C20	fullerenes	240	1.86	0.23	4659	60	-138
C36*	fullerenes	432	2.68	0.40	3729	108	-108
C60*	fullerenes	721	3.81	0.65	3186	180	-80
C70*	fullerenes	841	4.33	0.75	3098	210	-96
C240	fullerenes	2883	13.13	2.54	2742	720	-87
C20@C60	fullerenes	961	4.34	0.82	2720	240	-100
C20@C60@C240	fullerenes	3844	10.28	3.05	1611	960	-70
SCNT(10,0)	CNT	24164	108.58	21.86	2706	6010	-185
SCNT(6,6)*	CNT	25223	113.15	22.82	2701	6288	-153
SCNT(28,0)	CNT	70292	314.02	62.89	2690	1098	-489
DCNT(10,0)	CNT	29768	65.38	22.95	1323	1495	-245
DCNT(6,6)*	CNT	27148	61.19	21.37	1357	1454	-262
TCNT(10,0)	CNT	28070	50.58	21.94	1085	2173	-299
NR(6,6)	CNT	28266	92.50	23.58	1971	3204	-195
SCNT(16,0)@C60	CNT	27416	109.46	22.42	2404	1091	-448
MG	graphene sheets	26764	124.22	23.03	2795	6012	-143
BG	graphene sheets	27677.57	77.18	22.65	1679	6156	-142

In the original publication, orthogonal partial least square (OPLS) regression was used to develop nanoQSAR models to predict the interactions between the CNPs and the fragment. Three models were established and for each specific descriptors that best correlated the target E<sub>int</sub> were identified and used:

- The fullerene model used SSA as a descriptor ( $R^2=0.80$ , RMSE=0.485).
- The combined CNT and graphene sheets model used OSA and SDeg as descriptors ( $R^2=0.85$ , RMSE=0.440).
- The model for all three CNP family types used MW and SDeg as descriptors ( $R^2=0.80$ , RMSE=0.473).

However, the model validation process was not well explained (internal or external validation) so the reported statistics may be misleading. Due to the lack of information, we assume that the models are trained and tested on the same data.

### 7.1.2 Results of the first case study

Here we describe in detail the results of the three read-across modelling frameworks using the same data.

#### 7.1.2.1 EnaloskNN read-across model

This model aimed to quantify the relationship between physicochemical properties of the CNPs and their interaction energies with the cov-RNA fragment. The strength of the interaction may correlate with more stable adsorption of the cov-RNA fragment on the CNPs. This model was developed using KNIME, an object-oriented, open-source data integration, processing, and analysis platform (see Figure 30). Each node performs a specific task such as receiving data from a CSV file, filtering and transforming them, training machine learning algorithms, validating models, and visualising results. The Enalos+ extension nodes built on the existing KNIME infrastructure were used.

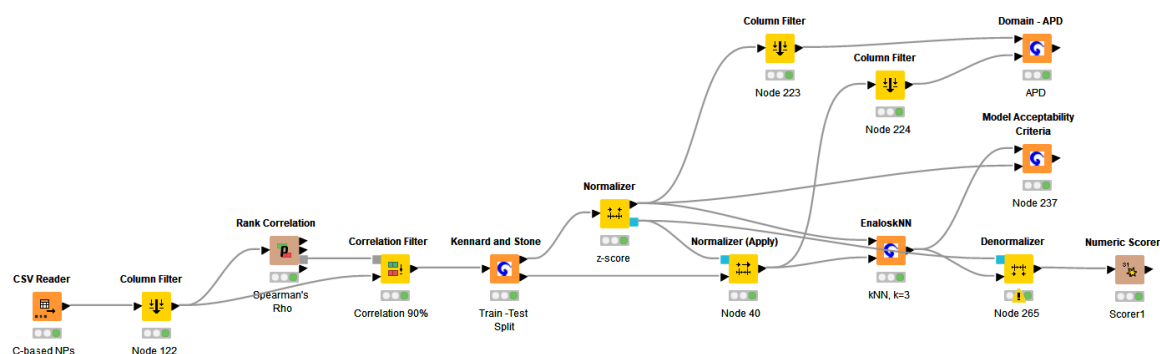


Figure 30: Model development workflow in KNIME Analytics platform using the Enalos+ nodes.

#### 7.1.2.1.1 Data pre-processing

Prior to modelling, it is essential to pre-process the data to transform, clean, and filter them. Feature selection was then applied before modelling to remove less relevant variables, avoid overfitting, and improve the models' predictive performance. Descriptors were assessed for correlations between them and with the target variable ( $E_{int}$ ). As the data were not normally distributed, Spearman's rank correlation coefficient (Spearman's rho) was used to assess the sign and magnitude of contribution descriptors made by calculating the difference in their rank. The coefficient can range between -1, indicating a strong negative correlation and 1, indicating a strong positive correlation, with 0 denoting no correlation. Descriptors that were correlated more than 90% with each other were removed. Rank correlation and correlation filter nodes in KNIME were used for variable selection. This resulted in the descriptor "Volume" being excluded from the model.

Because the remaining descriptors spanned different ranges, z-score (Gaussian) normalisation was implemented so they laid on a similar scale and were normally distributed. This involves subtracting the descriptor mean from each descriptor and dividing by the standard deviation of each descriptor set. This creates a new set of values whose mean is zero and standard deviation is one. The dependent variable ( $E_{int}$ ) was not normalised.

The dataset was divided into training and test sets in 70:30 proportion. The Kennard-Stone algorithm<sup>207</sup>, which creates the two subsets that each contain representative NP samples spanning the range of the original set, was used for partitioning. The model was trained on the training set and its performance assessed using the test set.

#### 7.1.2.1.2 Modelling methodology

*k*-nearest neighbours (*k*NN), a type of instance-based learning technique for regression tasks, was used to generate the model. *k*NN regression estimates the endpoint value of a new sample by averaging the endpoint values of the *k*-nearest samples in the training set. The distance between the new sample and all the samples in the training set is computed using a metric such as Euclidean distance. Next, the *k*-nearest samples are selected based on the computed distance, and the resulting average of their endpoint values is used as the predicted value of the endpoint target property for the new sample. The optimum number of nearest neighbours was defined as *k*=3 and the inverse distance used as the weighting factor for the nearest *k* points. Given its enriched read-across functionalities, the Enalos+ *k*NN KNIME node was used. This displays the nearest neighbours in the training set and the calculated distances for each instance in the test set. The node allows definition of the overall NPs neighbouring space, as recommended by read-across frameworks.

#### 7.1.2.1.3 Model validation

Model validation, an essential step in a read-across/nanoQSAR studies, provides evidence on a model's reliability and acceptability. For external validation, a simple technique of train-test split was implemented; the compounds in the training set were used for model development and those in the test set were used for evaluating the models' performance. Model prediction quality and robustness was validated by a several statistical measures such as the squared correlation coefficient ( $r^2$ ) between the predicted and observed energies, and dispersion-based metrics such as the MAE and the RMSE (Eqs 10, 3 and 5, respectively).

#### 7.1.2.1.4 Applicability domain

The DoA of a model is defined by the ranges of each descriptor used to generate the model. It defines the region of input data for which the model makes accurate predictions, prediction outside of the domain will be less accurate. The degree of similarity between the compound of interest and the model training set is determined by a distance-based method. Distance-based DoA assumes that the predictions obtained from a model are more likely to be valid for data points that are similar to the training data based on the Euclidean distances among them. The predefined applicability domain threshold is calculated according to Eq. 21.

Predictions where the distance from an external NM is smaller than the APD threshold value are considered reliable. The optimum threshold value was 1.96 and all predictions of the testing NPs fell inside the applicability domain (100% reliable predictions).

#### 7.1.2.1.5 Results

The  $k$ NN model was developed using the training set CNPs and the CNPs in the test set were used to assess its predictive power. In Table 11, the  $k=3$  nearest neighbours in the feature space and the distance from each one is presented, along with the actual and predicted energy values (see also Figure 31).

Internal and external model validation was implemented using the OECD validation principles. The statistical measures<sup>g</sup> were calculated for the external test set ( $r^2=0.95$ , MAE=15.4, RMSE=20.4 kJ/mol). Fitter models with a higher predictive power have a coefficient of determination close to 1, and small dispersion-based metrics. Golbraikh and Tropsha's test<sup>208</sup> for quality of fit and predictive ability of a model (see Appendix VI) was also applied, using the Enalos+ node for "Model Acceptability Criteria" (Figure 32).

The combined model has slightly better validation performance than the three models in the original publication. Nonetheless, direct comparisons are misleading, as the modelling methods used by the original publication are unclear.

To further assess the robustness of the model, Y-randomisation was applied<sup>89</sup>. As presented in Table 12, the models developed with scrambled endpoint values demonstrated lower performance thus, the original model is considered robust and valid.

---

<sup>g</sup> The respective formulas are presented in Appendix VI.



Table 11: The three closest neighbours from the training set and  $E_{\text{int}}$  predictions of test set CNPs.

Testing CNP	Neighbour 0	Distance N0	Neighbour 1	Distance N1	Neighbour 2	Distance N2	$E_{\text{int}}$ [kJ/mol]	Prediction [kJ/mol]
C36	C20	0.260	C20@C60	0.283	C240	0.298	-109	-110
C60	C20@C60	0.131	C240	0.158	C20	0.413	-80	-101
C70	C20@C60	0.106	C240	0.136	C20@C60@C240	0.436	-96	-92
SCNT(6,6)	SCNT(10,0)	0.050	NR(6,6)	0.551	TCNT(10,0)	0.838	-153	-192
DCNT(6,6)	DCNT(10,0)	0.041	TCNT(10,0)	0.145	SCNT(16,0)@C60	0.337	-262	-274

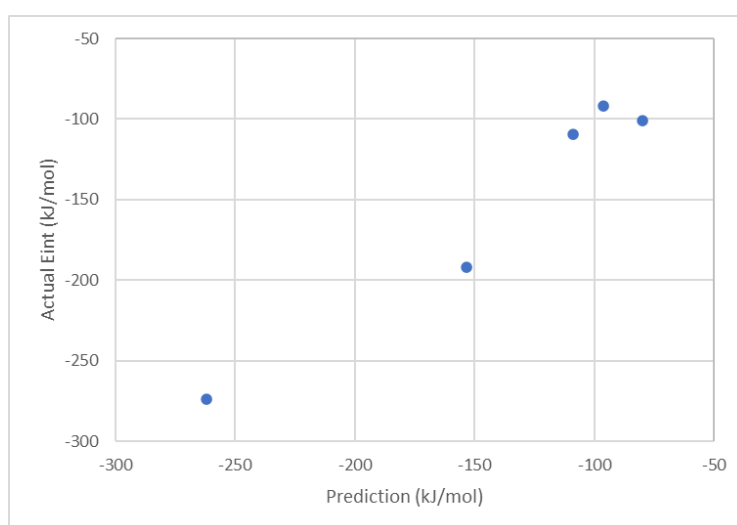


Figure 31: Actual vs predicted  $E_{\text{int}}$  values using the EnaloskNN method for the test set.

Criterion	Assessment	Result
$R^2 > 0.6$	PASS	$R^2 = 0.953$
$R_{\text{cvext}}^2 > 0.5$	PASS	$R_{\text{cvext}}^2 = 0.956$
$(R^2 - R_0^2)/R^2 < 0.1$	PASS	$(R^2 - R_0^2)/R^2 = 0.002$
$(R^2 - R'^2)/R^2 < 0.1$	PASS	$(R^2 - R'^2)/R^2 = 0.0$
$\text{abs}(R_0^2 - R'^2) < 0.3$	PASS	$\text{abs}(R_0^2 - R'^2) = 0.002$
$0.85 < k < 1.15$	PASS	$k = 0.913$
$0.85 < k' < 1.15$	PASS	$k' = 1.086$
<b>Model Predictive</b>		

Figure 32: Results of the “Model Acceptability Criteria” node using as input the EnaloskNN model.

Table 12: Statistic metrics on the test set for different random shuffles in a Y-randomization test repeated five times.

Y-randomisation	$r^2$	RMSE [kJ/mol]
1	0.001	153
2	0.004	73
3	0.603	121
4	0.266	100
5	0.330	97

Apart from having adequate predictive performance, a QSAR model should help elucidate mechanistic interpretations. Here, we discuss the influence of selected descriptors on the endpoint. Firstly, the “MW” is strongly correlated with the size of a molecule. The “structure” of each CNP account for the observation by Zhang *et al.* that the interaction affinity of the cov-RNA fragment increases in the order of fullerenes < graphene sheets < CNTs. The “SDeg” is a topological descriptor that represents the total number of degrees of non-hydrogen atoms connected to every atom of the CNP<sup>209</sup>. Surface area descriptors are important because nanostructures, fullerenes, CNTs and graphene have high surface to volume ratios that enable good adsorption and desorption<sup>210,211</sup>. The larger the surface area of the NM, the greater the number of adsorption sites available for a molecule to bind to. Finally, Zhang *et al.*<sup>127</sup> report that “Volume” correlated with SSA, which explains its high correlation with other variables and its eventual exclusion from the dataset.

A standardised QMRF report was created to document the model, where key information regarding the development, assessment and applicability of the model can be found (see Appendix V).

#### 7.1.2.2 Apellis model

The same modelling process was followed next using the rest of the tools. Apellis is a web-application for read-across model development (<https://apellis.jagpot.org/>) for NMs based on mathematical optimisation. During training, the application selects the most important NM properties that affect the endpoint to be modelled. Grouping NMs for performing read-across predictions involves the use of more than one similarity criterion (thresholds). The workflow applies a genetic algorithm to the grouping hypothesis (selected variables and threshold value) to generate the most accurate read-across estimations. For each query NM, neighbours are identified based on its distance from the training NMs and the threshold value. NMs with a distance equal or lower than the threshold value are considered neighbours for the query NM. Therefore, with this strategy, the number of neighbours is not fixed for all NMs due to local similarities that may lead to a dense or a spare neighbourhood. In fact, in cases of a strict threshold, no neighbours may be found for a query NM. The endpoint value is computed as the weighted average of the endpoint values of neighbour NMs.

The same dataset was used as input to train the model. However, an extra pre-processing step is essential before using the application. Input variables must be numeric so categorical variables need to be transformed into binary descriptors. As the “structure” values are categorical but not ordinal, the one-hot encoding technique was applied. Three binary variables were created, namely “Fullerene”, “CNT” and “Graphene”, and a “1” identifies the structure of each sample (Table 13).

Table 13: One-hot encoding for the categorical “structure” descriptor. For each label-value a binary variable was created.

CNPsa	Structure	Fullerene	CNT	Graphene
C20	fullerenes	1	0	0
C36	fullerenes	1	0	0
C60	fullerenes	1	0	0
C70	fullerenes	1	0	0
C240	fullerenes	1	0	0
C20@C60	fullerenes	1	0	0
C20@C60@C240	fullerenes	1	0	0
SCNT(10,0)	CNT	0	1	0
SCNT(6,6)	CNT	0	1	0
SCNT(28,0)	CNT	0	1	0
DCNT(10,0)	CNT	0	1	0
DCNT(6,6)	CNT	0	1	0
TCNT(10,0)	CNT	0	1	0
NR(6,6)	CNT	0	1	0
SCNT(16,0)@C60	CNT	0	1	0
MG	graphene sheets	0	0	1
BG	graphene sheets	0	0	1

The following parameters were tuned:

- Scaling of the raw data prior to modelling using the min-max normalisation,
- Dataset partitioning using the Kennard-Stone method and a training-test set ratio of 0.7:0.3,
- Number of chromosomes (candidate solutions) and number of generations (iterations) for the genetic algorithm, 20 and 50 respectively,
- Regularisation factor for the variable selection equal to 0 (no influence in the minimisation of the objective function).

The model was trained automatically, and the optimum similarity threshold was found to be 0.328 based on six selected variables: “MW”, “OSA”, the “volume”, “SSA”, “SDeg” and “Fullerene” binary variable. The resulting model was used to predict the properties of the test set, which produced the following validation statistics:

- $Q_{ext}^2 = 0.93$ ,
- RMSE = 25.5 kJ/mol,
- MAE = 20.6 kJ/mol.

The application plots neighbourhood space (neighbours of test samples in the training set) in a graphical format. The training results and the model were downloaded for further use. All the modelling results are presented in Figure 33.

All the test samples were within the DoA of the model as all samples had at least one neighbour in the training samples, allowing predictions to be calculated fairly accurately.

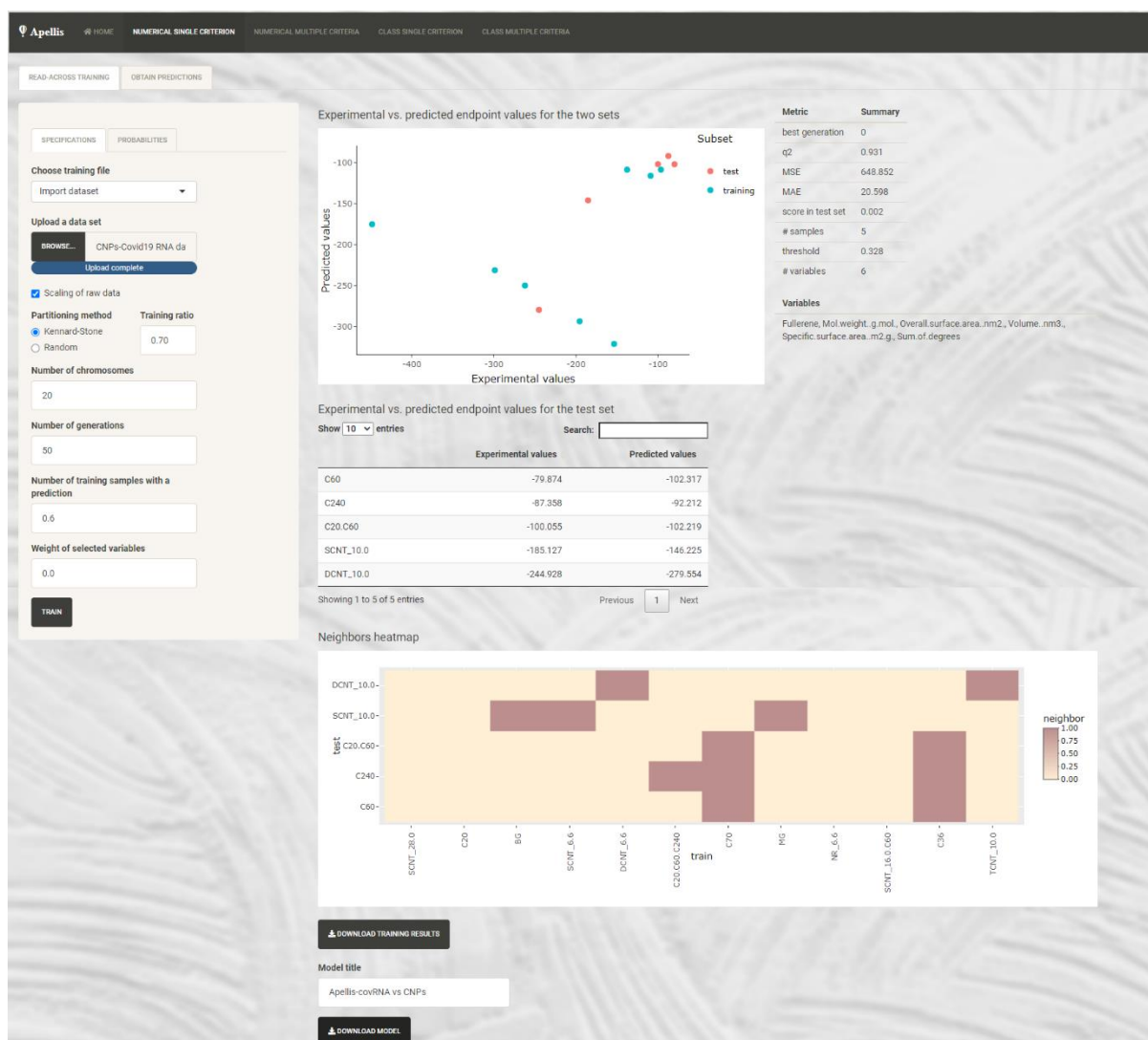


Figure 33: Screenshot of the modelling results using the Apellis web-application. The selected variables, the accuracy statistics, the predictions on the test set, and the plots for the actual vs predicted endpoint values, as well as the neighbouring space in heatmap format are presented.

### 7.1.2.3 Quantitative read-across (Read-Across-v4.1) model

The last assessed modelling tool is the Read-Across-v4.1 model. The Drug Theoretics and Cheminformatics Laboratory (DTC Lab) of the Jadavpur University (India) developed a quantitative read across software that predicts toxicity and biological activity of NMs or chemicals, using different similarity-based functions (Euclidean distance-based similarity, Gaussian kernel function similarity, Laplacian kernel function similarity)<sup>59</sup>. The software requires clean, ready-to-modelling data as training and test sets and tuning the model parameters including the similarity thresholds and the number of neighbours.

Data were prepared according to the requirements of the software and training and test sets defined as in the first model (EnaloskNN) presented in Table 10. One hot encoding was again applied to the data to transform the “structure” labels into binary values. The following hyper parameters were used:

- Gamma value for the Laplacian kernel of 1,
- Sigma value for the Gaussian kernel of 1,
- Number of similar training compounds of 3,
- Distance threshold of 0.5,

- Similarity threshold of 0.01.

The software generated in an excel file, predictions for each test sample, and accuracy statistics using the three similarity-based estimations (Figure 34).

ID	Yeuc(Test)	Ygk(Test)	Ylk(Test)
2	-108.514	-109.417	-113.051
3	-106.685	-100.193	-102.255
4	-86.9594	-90.768	-100.919
9	-253.412	-202.744	-205.466
12	-248.715	-253.308	-286.352
Q2f1=	0.767487	0.937543	0.918698
Q2f2=	0.490521	0.863145	0.821852
RMSEP=	46.84125	24.27705	27.69847
MAE=	29.97571	16.86741	21.47366

Figure 34: Screenshot of the output predictions and statistics using the DTC-Lab model, including one-hot encoding for the CNPs structures.

The procedure was repeated using the same parameters and input data but omitting the one hot encoding. The results are slightly improved as presented in Figure 35.

ID	Yeuc(Test)	Ygk(Test)	Ylk(Test)
2	-108.533	-109.417	-113.051
3	-106.504	-100.193	-102.255
4	-87.0732	-90.768	-100.919
9	-157.747	-159.399	-162.659
12	-248.883	-253.308	-286.352
Q2f1=	0.979111	0.988242	0.974194
Q2f2=	0.954228	0.974237	0.943455
RMSEP=	14.03996	10.53335	15.60499
MAE=	10.7464	8.198366	12.9123

Figure 35: Screenshot of the output predictions and statistics using the DTC-Lab model, without one-hot encoding for the CNPs structures.

DoA limits or reliability of the predictions were not provided within the tool.

### 7.1.3 Discussion

The results of the three methods are presented briefly in Table 14 and in Table 15. Direct comparisons can be made between the EnalosKNN model and the DTC-Lab models considering that they were trained and validated using the same train and test set splits. Even if the same splitting ratio and algorithm were used in the Apellis model, slight differences in the Kennard-Stone algorithm still led to different training and test sets. Conspicuously, there is no way of providing user-defined sets directly to the application for prediction. Thus, comparisons with the two other methods cannot be made. Prediction performance is best assessed by the MAE and RMSE statistics (MAE is less sensitive to 1-2 large outliers). We observe that all three methods produce quite similar validation statistics.

Considering that the EnaloskNN and DTC-Lab models are trained and tested on the same data, consensus predictions can be made to avoid any biases inserted by the models, by averaging the predictions of the individual models. The consensus model results presented in Table 14 **Error! Not a valid bookmark self-reference.** makes predictions of similar accuracy to those of the DTC-Lab Gaussian kernel model.

Table 14: Actual and predicted  $E_{\text{int}}$  values in kJ/mol, according to the EnaloskNN and DTC-Lab models (without the “structure” variables). Consensus predictions based on the average of the EnaloskNN and DTC-Lab predictions are also presented. Validation statistics are also presented.

CNP	Actual values [kJ/mol]	EnaloskNN [kJ/mol]	DTC-Lab: Euclidean kernel [kJ/mol]	DTC-Lab: Gaussian kernel [kJ/mol]	DTC-Lab: Laplacian kernel [kJ/mol]	Consensus [kJ/mol]
C36	-109	-110	-109	-109	-113	-110
C60	-80	-101	-107	-100	-102	-102
C70	-96	-92	-87	-91	-101	-93
SCNT(6,6)	-153	-192	-158	-159	-163	-168
DCNT(6,6)	-262	-274	-249	-253	-286	-266
	MAE	15.4	10.8	8.2	12.9	9.1
	RMSE	20.4	14.0	10.5	15.6	12.2
	$Q_{\text{ext}}^2$	0.96	0.98	0.99	0.97	0.98

Table 15: Actual and predicted  $E_{\text{int}}$  values in kJ/mol, according to the Apellis model. Validation statistics are also presented.

CNP	Actual values [kJ/mol]	Apellis predictions [kJ/mol]
C60	-80	-102
C240	-87	-92
C20@C60	-100	-102
SCNT(10,0)	-185	-146
DCNT(10,0)	-245	-280
	MAE [kJ/mol]	20.6
	RMSE [kJ/mol]	25.5
	$Q_{\text{ext}}^2$	0.93

The EnaloskNN model, although less accurate, is easier to use because of its integration within the KNIME Analytics platform and user-friendly environment. Stakeholders can create visual data flows consisting of nodes and connections and they have flexibility to exploit all the KNIME features and tools for data pre- and post-processing, modelling, validation, and visualisation of the results. Therefore, users can have complete oversight of the analysis workflow and can tune the parameters at each step. The EnaloskNN node itself, apart from the  $k$ NN prediction, allows the inspection of the neighbouring space of NMs, a pre-requisite for the read-across framework. Nonetheless, the Enalos+ nodes are proprietary software and can be only used under a paid licence scheme.

The Apellis web application also allows the read-across model development using a user-friendly interface. No coding is needed, and all parameter tuning is performed through menus and buttons. The background optimisation methods produce an automatic grouping hypothesis that leads to trustworthy, predictive results using one or two similarity criteria. In addition, the application performs variable selection, identifying the most informative descriptors for the endpoint predictions. Users can train and evaluate predictive read-across models using their own data through Apellis, then share the

generated models as ready-to-use web applications with the nano community. The Apellis tool comes with a variety of visualisation tools and services that display the results graphically in the form of simple-to-understand graphs and tables (predictions, neighbouring space). However, considering that the background method is based on a genetic algorithm, depending on the input data and parameters, training can take some time to find an optimised grouping hypothesis. This is in practice a minor issue as the training of the model is performed once, and it can be used several times to perform almost instantaneous predictions. However, in hand, delays in algorithm convergence led several times to server timeouts and interruptions of the training requiring training to be repeated from scratch. Another problem was the lack of flexibility in the splitting of the input data (there is no option to provide user-defined training and test sets), in the use of categorical-labelled variables as descriptors (one-hot encoding had to be performed). Additionally, any additional robustness tests must be applied using other software. Nonetheless, the available functionalities through the application may be considered satisfactory for a non-informatics expert.

Finally, the DTC-Lab software is a powerful tool for read-across predictions and exploration of the neighbouring space. However, these functionalities cannot be exploited easily as the software is complex and not very user-friendly. Since the software is executed serially through consecutive dialog boxes, if an invalid parameter value or the wrong file is introduced, the program terminates, and all procedures must be carried out again from the beginning. Data input format took time to understand as the training material was not comprehensive. An improved interface would allow for simultaneous display of all fields and the adjustment of the parameters in the case of an error, or in case of repetitive execution in a sensitivity analysis scheme. In addition, many of the results presented in the output are not explained. A more detailed user guide could better motivate the stakeholders to use the tool. In addition, all pre-processing steps including variable selection, one-hot encoding, and dataset splitting should be performed prior to using the tool, and any post-processing robustness tests should be performed using other software. Finally, the DoA or the reliability of the predictions is not provided within the tool.

## 7.2 Case study 2: Assessment of different tools for the prediction of TiO<sub>2</sub>-based NMs' endpoints

### 7.2.1 Scope of the second case study

The global NP titanium dioxide market was valued at €11.1 billion in 2020 (based on the average Euro – USD 2020 exchange rate), and is projected to reach €18.7 billion by 2030 based on the current (May 2023) average Euro – USD exchange rate<sup>212</sup>. Large quantities of TiO<sub>2</sub> NPs are used in coatings, inks and paints, suncreams, toothpastes, food colorants and pharmaceuticals, etc.<sup>213–215</sup> Their band gap energies (3.2eV for anatase TiO<sub>2</sub>) means they absorb ultraviolet (UV) light<sup>216</sup>. To maximise absorption of visible light (e.g., for photocatalytic properties under sunlight) doping of TiO<sub>2</sub> NPs with organic and inorganic materials reduces the activation energy and narrows the bandgaps<sup>215,216</sup>.

TiO<sub>2</sub> NPs have different physicochemical characteristics from their larger particle forms that alter their properties. The most significant negative effects of exposure to TiO<sub>2</sub> NPs in experimental animals are lung cancer and pulmonary inflammatory reactions. Inhalation and dermal exposure to TiO<sub>2</sub> NPs are typically the main methods of entry into the body during work exposure or product use<sup>213</sup>. TiO<sub>2</sub> NPs in personal care and household products are released to sewage and enter wastewater treatment plants. They impact coastal ecosystems that support fishing and leisure activities as well as phytoplankton and they prevent algae growth by altering membrane structure through elevated lipid peroxidation<sup>214</sup>.

The need to assess NM properties and adverse effects before time and resource intensive synthesis and testing has spurred the development of computational models that can identify undesirable

candidate NMs. In fact, techniques and tools for the computational evaluation of NMs adverse effects and properties have grown considerably in recent years. However, because of lack of data science, statistics, or programming skills, nanosafety researchers who could benefit from these *in silico* methodologies have not adopted them to a large extent. To address this need, several models are available through a GUI as online apps. Using a GUI designed specifically for non-informatics specialists, the various technical aspects of the models are largely hidden, with only elements required for running the models being displayed. This allows a wider range of scientists to use computational models to design novel NMs design and for risk evaluation. The diverse web-tools available for prediction of TiO<sub>2</sub> NMs endpoints are summarized in the literature review section. The web-tools assessed in this case study are:

- The QSAR for nano-mixtures tool ([https://kriktcsrc.shinyapps.io/TiO2\\_Dmagna/](https://kriktcsrc.shinyapps.io/TiO2_Dmagna/)) for the prediction of toxicological endpoints of TiO<sub>2</sub>-based nano-mixtures<sup>130</sup>.
- The NanoMixHamster tool (<https://nanomixhamster.cloud.nanosolveit.eu/>) for the cytotoxicity prediction of TiO<sub>2</sub>-based multicomponent NMs toward Chinese hamster ovary (CHO-K1) cells<sup>101</sup>.
- The NanoToxRadar tool (<https://nanotoxradar.kitox.re.kr/>) for the cytotoxicity to A549 cell line and the zeta potential predictions of metal, MeOx and polystyrene cored NMs<sup>201</sup>. The tool was employed for TiO<sub>2</sub>-based NMs.
- The INSIDE nano tool (<http://inano.biobyte.de/>) for the reveal of connections between phenotypic entities based on their effects on genes<sup>167</sup>. The functionalities of the tools are demonstrated for TiO<sub>2</sub> NMs.
- The lung exposure dose calculator (<https://lungexposure.cloud.nanosolveit.eu/>) for the estimation of acute regional lung-deposited dose of inhaled inorganic engineered NMs (TiO<sub>2</sub> in this case) in humans. The use of the tool is combined with the multi-box aerosol model (<https://aerosol.cloud.nanosolveit.eu/>) for the assessment of human exposure to NMs in indoor environments<sup>161</sup>.
- The IATA: integration of physiologically based pharmacokinetic (PBPK) model and occupational exposure tool for the simulation of NM biodistribution (TiO<sub>2</sub> in this case) in humans due to inhalation (<https://exposurepbpk.cloud.nanosolveit.eu/>). The use of the tool is combined with the multi-box aerosol model (<https://aerosol.cloud.nanosolveit.eu/>) for the assessment of human exposure to NMs in indoor environments<sup>161</sup>.

## 7.2.2 Results of the second case study

### 7.2.2.1 QSAR for nano-mixtures

This web application ([https://kriktcsrc.shinyapps.io/TiO2\\_Dmagna/](https://kriktcsrc.shinyapps.io/TiO2_Dmagna/)) generates the predictions of two ecotoxicity endpoints: EC<sub>50mix</sub> and immobilisation (percentage of non-mobile and dead water fleas relative to control) in *D. magna* exposed to TiO<sub>2</sub>-based nano-mixtures<sup>130</sup> (Figure 36). The required input data are entered through dropdown menus and sliders. Depending on the type of mixture descriptors selected, different random forest models are employed to calculate the ecotoxicity endpoints. Model performance statistics is presented on the interface. In both cases, the predicted values for EC<sub>50mix</sub> or for the immobilisation (%) percentage of the nano-mixture are automatically presented along with indication of the ecotoxicity risk level in a colour-coded format (green- “low”, orange- “medium”, and red- “high”). The application can be used for a toxicity and safety limits assessment by studying different combinations of concentrations and mixtures, to reduce the cost of the environmental risk assessment. For each toxicity endpoint, we explored the predicted endpoint values as function of model input parameters range.



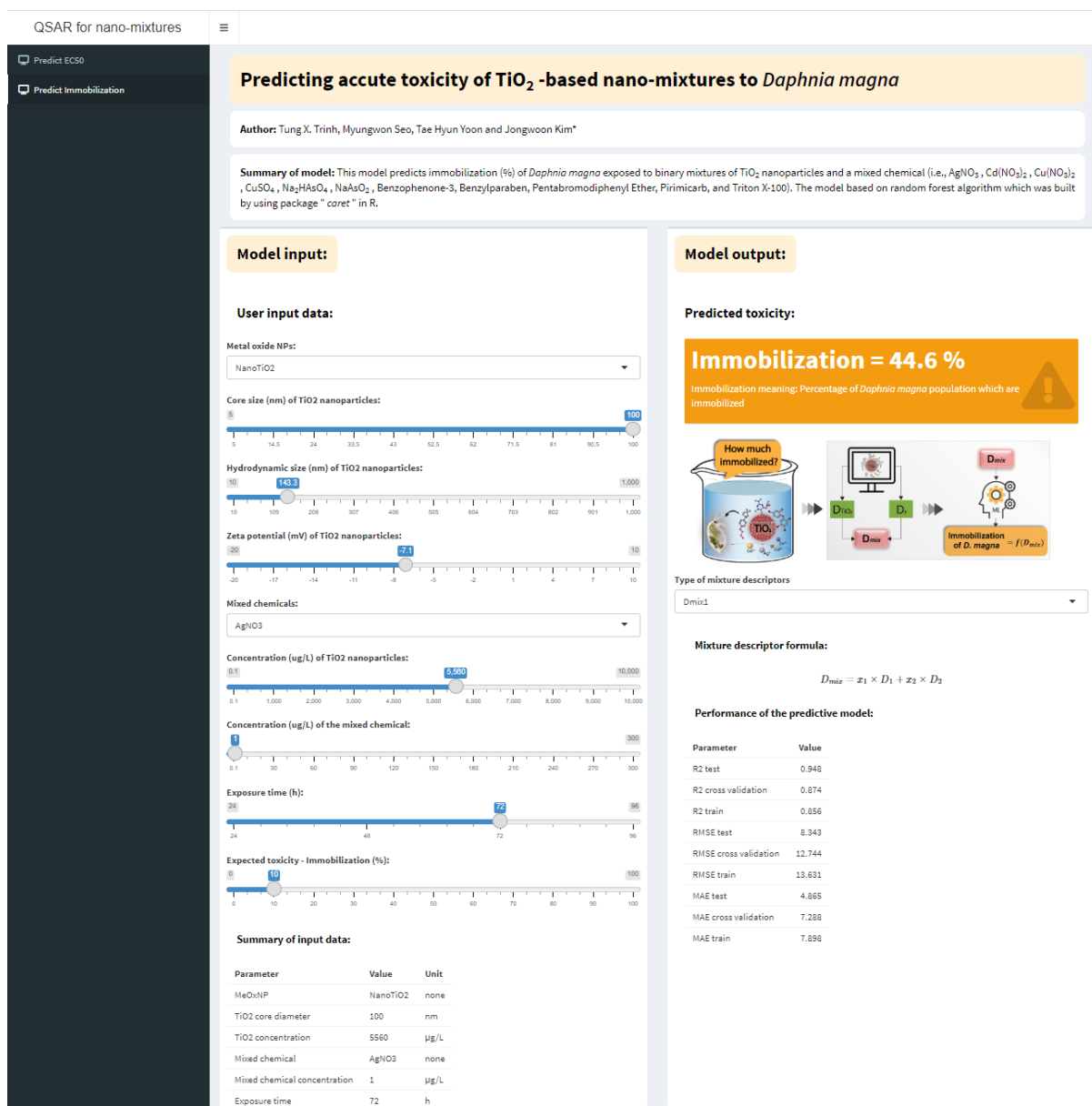


Figure 36: Screenshot of the user interface of the QSAR for nano-mixtures-tool. The users can tune the input parameters from the left-hand side of the interface and observe the predicted toxicity endpoint in the right-hand side of the interface according to the selected type of mixture descriptors. The performance statistics of the model and a summary of the input parameters are also presented within the application. The same applies for both the predicted endpoints (immobilisation and  $EC_{50mix}$ ).

#### 7.2.2.1.1 Immobilisation prediction

To predict the (%) immobilisation of *D. magna* (acute toxicity) of the  $TiO_2$  based nano-mixtures, users input the core and hydrodynamic size of the  $TiO_2$  NPs, zeta potential values, the type of mixed chemical, its concentration (mass fraction), the type of mixture descriptors, and the exposure time of the nano-mixtures. However, when choosing a different mixed chemical (not the default  $AgNO_3$ ), the model generates the same predictions and summary i.e., the same results were produced for all chemicals. Furthermore, altering the input core, hydrodynamic sizes, and zeta potential values should not be performed randomly, as these values are dependent on one another<sup>142</sup> (e.g., the NPs hydrodynamic size and zeta potential values depend on NPs core size). In absence of experimental data, and to remain in the applicability domain of the model, we used the data from the supplementary information file

from Trinh *et al.*<sup>130</sup> (core diameter of 100 nm, hydrodynamic diameter of 143.3 nm<sup>h</sup>, zeta potential of -7.1 mV, concentration of TiO<sub>2</sub> NPs of 5560 µg/L, and concentration of the mixed chemical of 1 µg/L). Next, we assessed the immobilisation predictions as a function of the exposure time (24, 48, 72 and 96 hours) for the two available mixture descriptors (arithmetic mean and geometric mean-based). As Figure 37 shows, longer exposure generate higher toxicity than in lower exposure times, in accord with intuition.

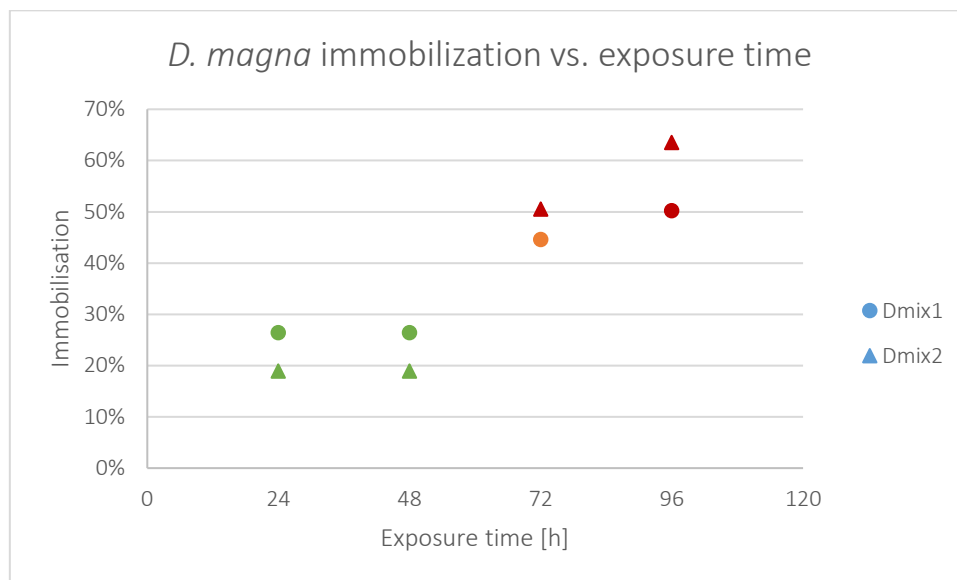


Figure 37: Immobilisation (%) predicted values for AgNO<sub>3</sub>-TiO<sub>2</sub> nano-mixtures, as a function of the exposure time, using the D<sub>mix1</sub> (arithmetic mean-based) and the D<sub>mix2</sub> (geometric-mean based) descriptors. Markers in green correspond to low risk levels, in orange to medium risk levels and in red in high risk levels, as generated by the web application.

#### 7.2.2.1.2 EC<sub>50mix</sub> prediction

To predict the half-maximal effective concentration (EC<sub>50mix</sub>) of a mixture towards *D. magna* at 48 hours, users should select from the dropdown menus the mixed chemical and the concentration fraction of the TiO<sub>2</sub> NPs in the nano-mixture. Figure 38 and Figure 39 summarize the predictions for the combination of different mixed chemicals in different fractions (TiO<sub>2</sub> mole fractions of 0.05, 0.15, 0.3, 0.5, 0.75, and 0.9) and for different mixture descriptors (arithmetic mean and geometric mean-based). In both cases, the nano-mixtures are safer (high EC<sub>50mix</sub> values) for lower TiO<sub>2</sub> concentrations in the mixtures. We also observe that when using the geometric mean-based descriptors (D<sub>mix2</sub> in the application), the EC<sub>50mix</sub> values using the same fractions are higher (less toxic nano-mixtures). However, bear in mind that the predictions made using this package are mixture compound agnostic. According to the corresponding publication of Trinh *et al.*<sup>130</sup>, the use of geometric mean-based descriptors (D<sub>mix8</sub> in the publication) leads to better performing models thus, the predictions of Figure 39 should be considered more accurate.

<sup>h</sup> The hydrodynamic diameter values were not available through the original publication; thus, the value was retrieved from the publication of Park *et al.*<sup>254</sup> to which Trinh *et al.*<sup>130</sup> refer to.

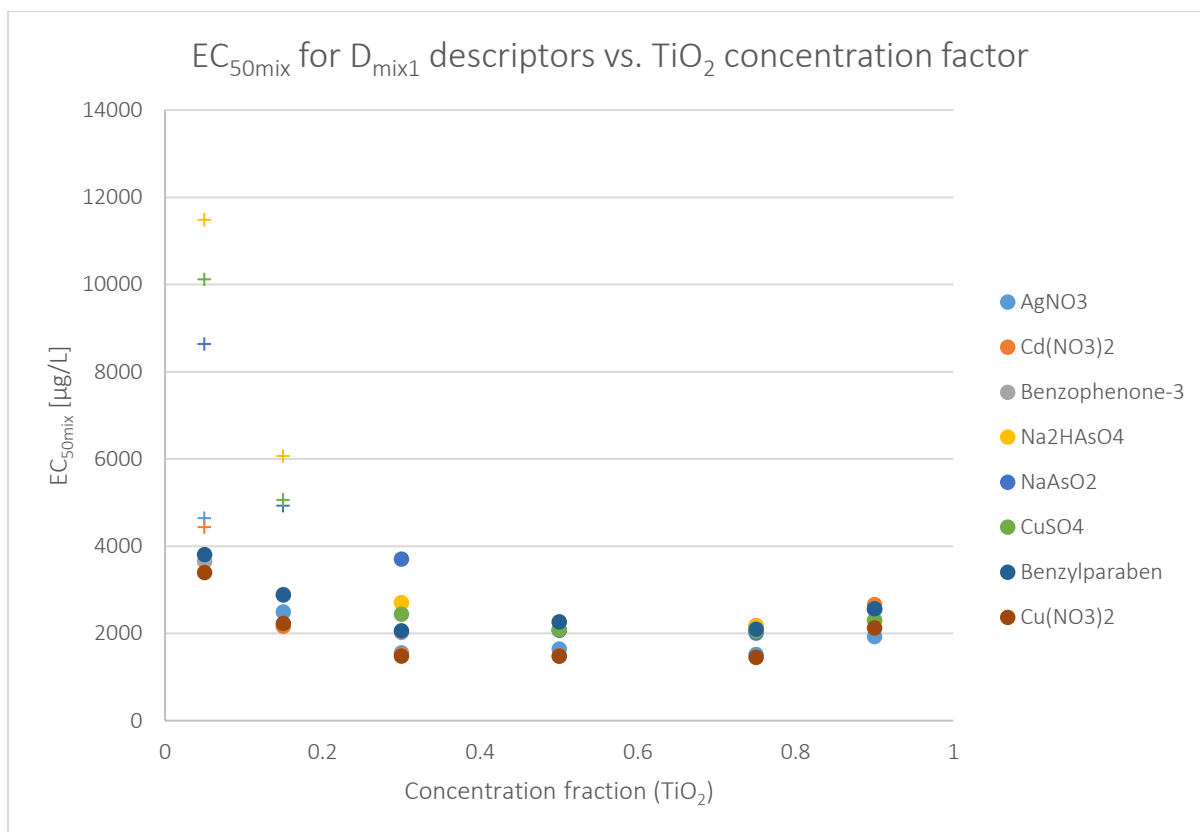


Figure 38: EC<sub>50mix</sub> predicted values for different TiO<sub>2</sub> based nano-mixtures, as a function of the mole fraction of the TiO<sub>2</sub> into the mixture, using the D<sub>mix1</sub> (arithmetic mean-based) descriptors. Markers represented as crosses correspond to low risk levels, whereas circular markers correspond to medium risk levels, as generated by the web application.

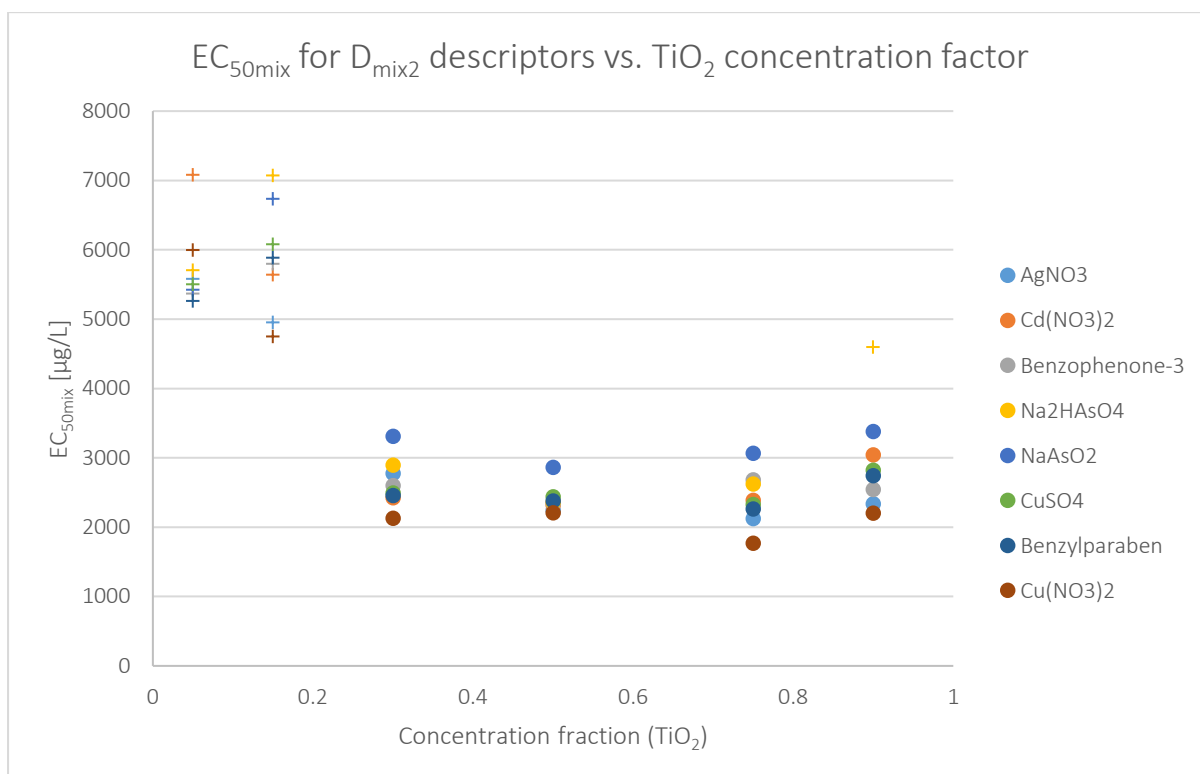


Figure 39: EC<sub>50mix</sub> predicted values for different TiO<sub>2</sub> based nano-mixtures, as a function of the mole fraction of the TiO<sub>2</sub> into the mixture, using the D<sub>mix2</sub> (geometric-mean based) descriptors. Markers represented as crosses correspond to low risk levels, whereas circular markers correspond to medium risk levels, as generated by the web application.

### 7.2.2.1.3 Discussion

Overall, the application can be used to assess the safety limits of the nano-mixtures or to predict the (%) immobilisation of *D. magna* when experimental NP feature data are available. The intuitive environment allows users to make the toxicity-related endpoint predictions in a short time with minimal training. However, improvements are needed to expand the use of the tool by the stakeholders:

- Although the application is briefly described in the relevant publication of Trinh *et al.*<sup>130</sup>, a more comprehensive manual is needed for stakeholders, including a detailed explanation of the input values. For example, the meaning of the two sliders “Expected toxicity - Immobilization (%)” and “Expected EC<sub>50mix</sub> (ug/L)” in the two tabs is not obvious, and changing these parameters did not affect the results. The colour-coded ecotoxicity risk level limits (low/medium/high) are also not explained or linked to a specific regulatory framework, guideline, or other literature source.
- There are inconsistencies between the models described in the publication and those in the web application such as:
  - In the publication, mixture descriptors are calculated using eight different formulae ( $D_{mix1}$ -  $D_{mix8}$ ) however, only two types of mixture descriptors are available within the tool ( $D_{mix1}$  and  $D_{mix2}$ ). Although this may not be an issue, a better explanation is needed (e.g., in a training manual) to avoid misunderstandings in the proper use of the application. In Figure 40 the differences between the publication and the application are presented.

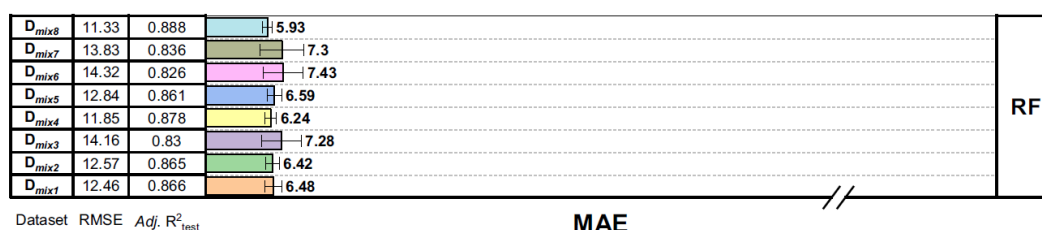
A	B
$D_{mix1} = x_1D_1 + x_2D_2$	<div>Type of mixture descriptors</div> <div>Dmix1</div>
$D_{mix2} = (x_1D_1 + x_2D_2)^2$	<div>Mixture descriptor formula:</div> $D_{mix} = x_1 \times D_1 + x_2 \times D_2$
$D_{mix3} = \sqrt{(x_1D_1)^2 + (x_2D_2)^2}$	<div>Type of mixture descriptors</div> <div>Dmix2</div>
$D_{mix4} = \sqrt{x_1}D_1 + \sqrt{x_2}D_2$	<div>Mixture descriptor formula:</div> $D_{mix} = \sqrt{(x_1 \times D_1 \times x_2 \times D_2)}$
$D_{mix5} = x_1D_1^2 + x_2D_2^2$	
$D_{mix6} = x_1^2D_1 + x_2^2D_2$	
$D_{mix7} = \sqrt[3]{x_1^3D_1 + x_2^3D_2}$	
$D_{mix8} = \sqrt{(x_1 \times  D_1 ) \times (x_2 \times  D_2 )}$	

Figure 40: Mixture-descriptors as presented [A] in the publication of Trinh *et al.*<sup>130</sup> and [B] in the relevant application. The  $D_{mix2}$  formula of the application, corresponds to the  $D_{mix8}$  formula of the publication.

- The DoA of the models is assessed in the relevant publication but not provided in the web application.
- There are inconsistencies in the use of the input descriptors for each model. For example, the publication states that “for models predicting  $\log EC_{50}$ , mixture descriptors using mass fraction produced higher performance models (lower MAE and RMSE, and higher  $adj. R^2_{test}$ ) than the mixture descriptors using mole fraction. In contrast, for models predicting immobilization, mixture descriptors using mole fraction gave better models than the mixture descriptors using mass fraction.” However, in the

application the mass fraction is used as input for the immobilisation prediction and the mole fraction is used for the EC<sub>50mix</sub> prediction.

- Conspicuously, the statistics in the application are not the same as in the publication, as shown by Figure 41 and Figure 42. Further explanation is necessary to prevent misinterpretations of the results (e.g., in the case of EC<sub>50mix</sub> it is probable that the model in the publication is trained using log transformed EC<sub>50mix</sub> values whereas the model in the application is trained using the original EC<sub>50mix</sub> values).



**Mixture descriptor formula:**

$$D_{mix} = x_1 \times D_1 + x_2 \times D_2$$

**Performance of the predictive model:**

Parameter	Value
R2 test	0.948
R2 cross validation	0.874
R2 train	0.856
RMSE test	8.343
RMSE cross validation	12.744
RMSE train	13.631
MAE test	4.865
MAE cross validation	7.288
MAE train	7.898

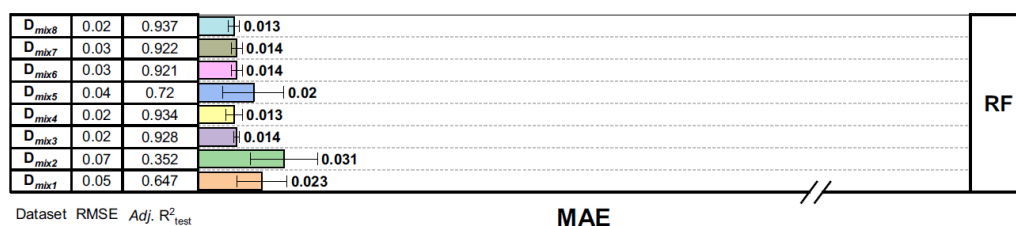
**Mixture descriptor formula:**

$$D_{mix} = \sqrt{(x_1 \times D_1 \times x_2 \times D_2)}$$

**Performance of the predictive model:**

Parameter	Value
R2 test	0.937
R2 cross validation	0.879
R2 train	0.861
RMSE test	10.721
RMSE cross validation	12.524
RMSE train	12.939
MAE test	6.389
MAE cross validation	7.130
MAE train	7.316

Figure 41: Model performance statistics for the immobilisation prediction as presented in the publication of Trinh *et al.*<sup>130</sup> (top) and in the relevant application (bottom).



Mixture descriptor formula:

$$D_{mix} = x_1 \times D_1 + x_2 \times D_2$$

Performance of the predictive model:

Parameter	Value
R2 test	0.951
R2 cross validation	0.906
R2 train	0.924
RMSE test	458.351
RMSE cross validation	729.944
RMSE train	786.689
MAE test	371.001
MAE cross validation	408.376
MAE train	418.342

Mixture descriptor formula:

$$D_{mix} = \sqrt{(x_1 \times D_1 \times x_2 \times D_2)}$$

Performance of the predictive model:

Parameter	Value
R2 test	0.962
R2 cross validation	0.869
R2 train	0.848
RMSE test	658.415
RMSE cross validation	936.563
RMSE train	997.725
MAE test	540.828
MAE cross validation	488.425
MAE train	474.451

Figure 42: Model performance statistics for the EC<sub>50</sub>mix prediction as presented in the publication of Trinh *et al.*<sup>130</sup> (top) and in the relevant application (bottom).

- Finally, as stated before, in the tab for immobilisation prediction the application produces the same results for all types of mixture chemicals.

#### 7.2.2.2 NanoMixHamster

The NanoMixHamster<sup>101</sup> (<https://nanomixhamster.cloud.nanosolveit.eu/>) is an application for predicting toxicity (pEC<sub>50</sub>) values of TiO<sub>2</sub>-based multicomponent NMs (Figure 43) against the adult Chinese hamster ovary (CHO-K1) cell line using the structure-activity prediction networks (SAPNet) methodology<sup>112</sup>. It consists of two main steps: the definition of the metallic NM composition (TiO<sub>2</sub> and Au, Ag, Pt, or Pd) to acquire its additive electronegativity value; and the toxicity prediction based on the results obtained in the previous step. The summary tab includes three graphs to visualize the data used to develop the model and the model applicability domain (Figure 44). Finally, the dataset generator tab can be used to generate a new dataset of theoretical TiO<sub>2</sub>-based multicomponent NMs with mixtures of silver, gold, palladium, and platinum. In a next step, it is possible to generate the predictions of the additive electronegativity and the cytotoxicity values for all the generated dataset simultaneously.



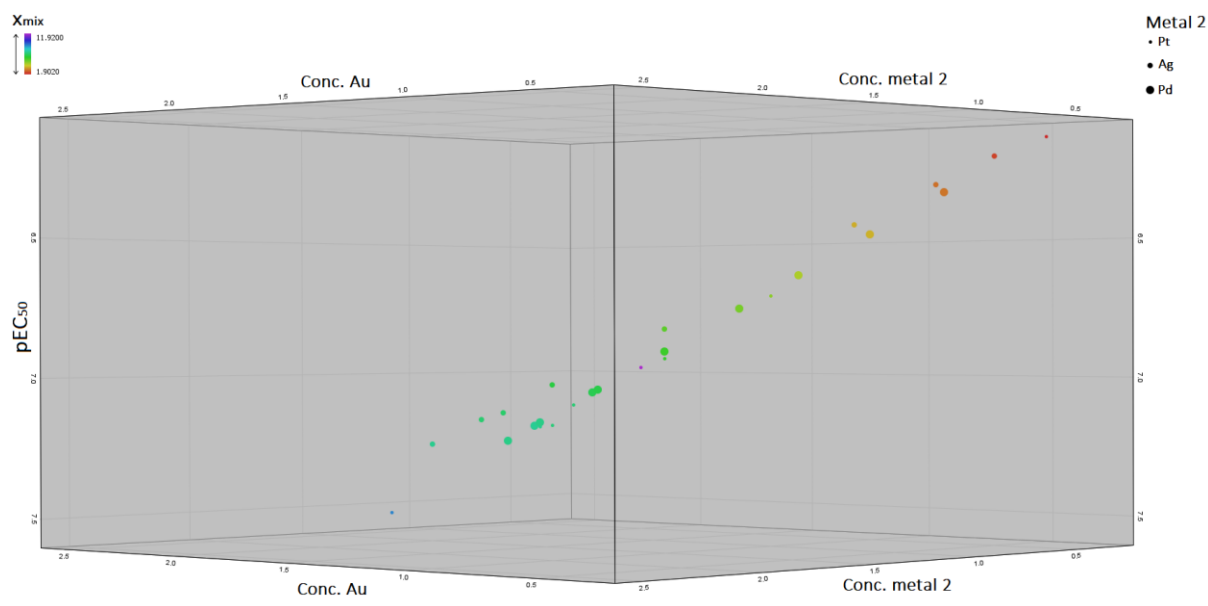


Figure 45: Graphical representation of the predictions of the cytotoxicity ( $pEC_{50}$ ) of  $TiO_2$  multicomponent NPs modified with Au and Ag, Pt and Pd (referred as metal 2 and encoded in the size of the markers). The additive electronegativity values ( $x_{mix}$ ) are encoded in the colour scale of the markers.

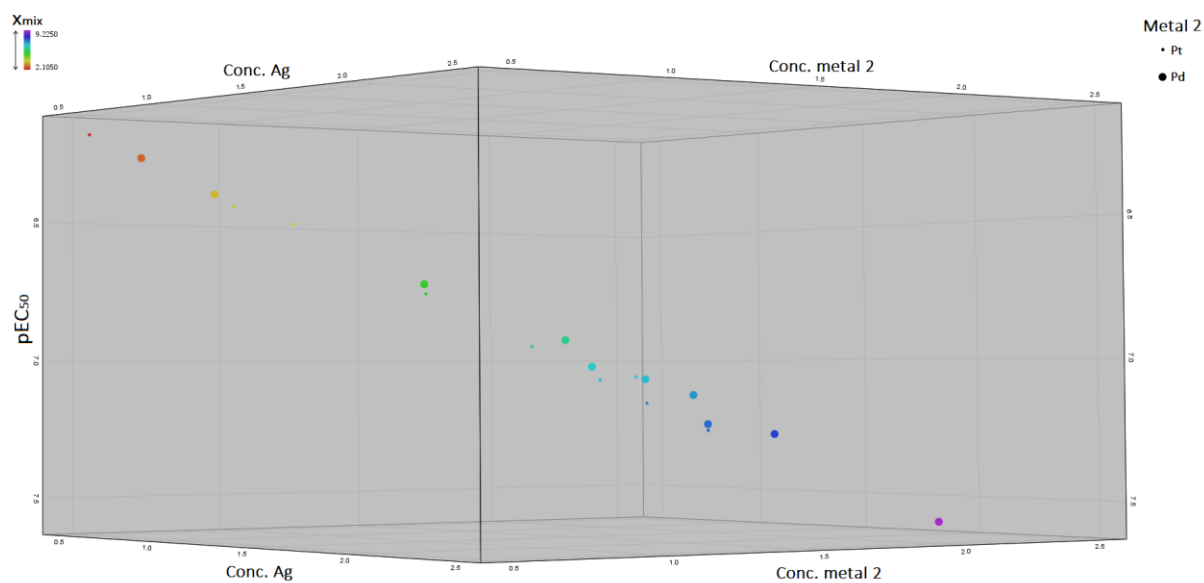


Figure 46: Graphical representation of the predictions of the cytotoxicity ( $pEC_{50}$ ) of  $TiO_2$  multicomponent NPs modified with Ag and Pt and Pd (referred as metal 2 and encoded in the size of the markers). The additive electronegativity values ( $x_{mix}$ ) are encoded in the colour scale of the markers.



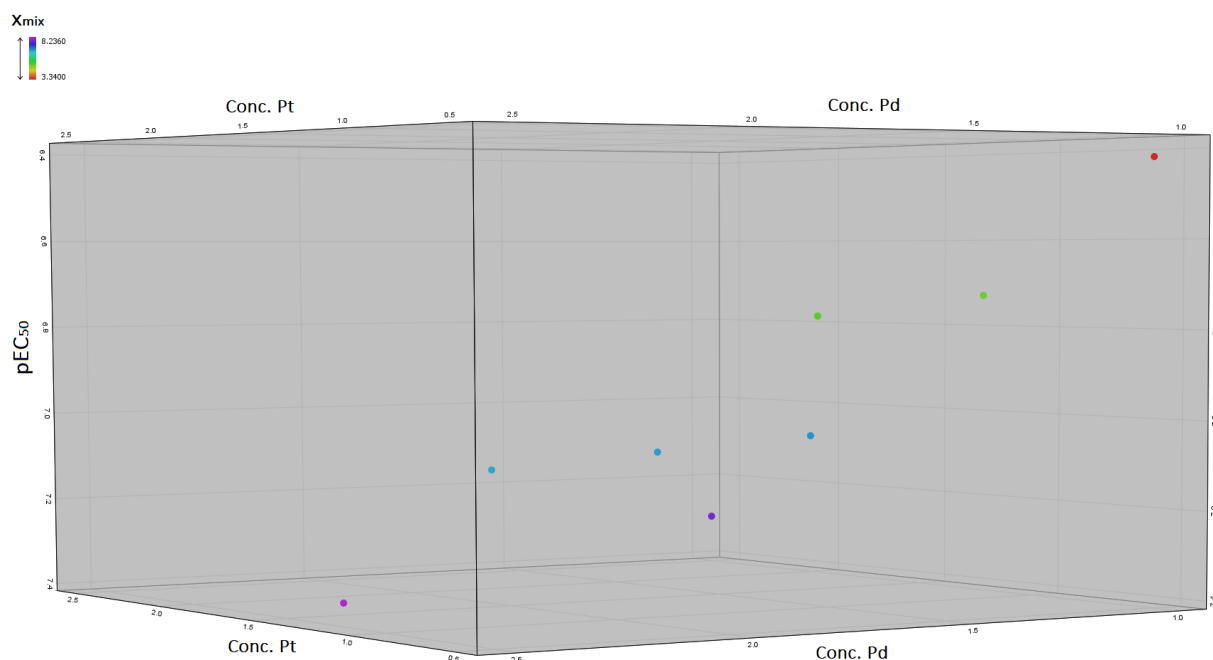


Figure 47: Graphical representation of the predictions of the cytotoxicity ( $pEC_{50}$ ) of  $TiO_2$  multicomponent NPs modified with Pt and Pd. The additive electronegativity values ( $x_{mix}$ ) are encoded in the colour scale of the markers.

#### 7.2.2.2.1 Discussion

The application is simple to use and quickly generates predictions for the additive electronegativity and the toxicity of  $TiO_2$ -based multicomponent NMs and reliability of the latter based on the domain of the applicability of the model. One advantage of the application is the ability to generate simultaneous predictions of the additive electronegativity and the cytotoxicity values of a dataset of hypothetical  $TiO_2$ -based multicomponent NMs modified with Au, Ag, Pt, or Pd nanoclusters. This feature allows screening of these NMs for adverse effects prior to their synthesis, saving time and resources during the design and the risk assessment of NMs. The application is described in detail in Stoliński *et al.*<sup>101</sup> thus, a training manual is not essential or possibly, required.

The SAPNet approach<sup>112</sup> on which the NanoMixHamster is based involves a series of interconnected predictive models. This combination of models is useful and necessary to exploit all tools and models available in the nanoinformatics field. Also, it makes possible the direct assessment of the influence of specific properties to the endpoint of interest (e.g., toxicity) through this row of consecutive models. Therefore, in a SbD scenario, input properties can be modified properly to ensure the final NM is safe and has the desired properties. Nonetheless, in combining models (especially if they are trained on different data) researchers should be cautious about combining properties (e.g., if they are measured under the different conditions or protocols in the combined models), about the reproduction and propagation of uncertainty through the network, and about the combined applicability domains of these models. These ambiguities will hinder stakeholders from trusting the final prediction and using it in their applications.

#### 7.2.2.3 NanoToxRadar

This tool (<https://nanotoxradar.kitox.re.kr/>, Figure 48) allows the use of the size-dependent electron configuration fingerprint model<sup>201</sup> to predict NMs cytotoxicity and zeta potential through a user-friendly interface. Users provide the NM's core composition (different metal, MeOx, polymeric and inorganic NMs), the doping parameters, coating materials, and size parameters such as NM shape (sphere or rod). The application predicts the cytotoxicity (more than 20% cell death of A549 cells at 100  $\mu$ g/mL)

probability value, the cytotoxicity benchmark dose (BMD5) value at a 5% benchmark response in a cytotoxicity assay using A549 cells, the zeta potential values, and an indication of the risk according to cytotoxicity probability (blue-“safe”, grey-“uncertain”, orange-“warning”).

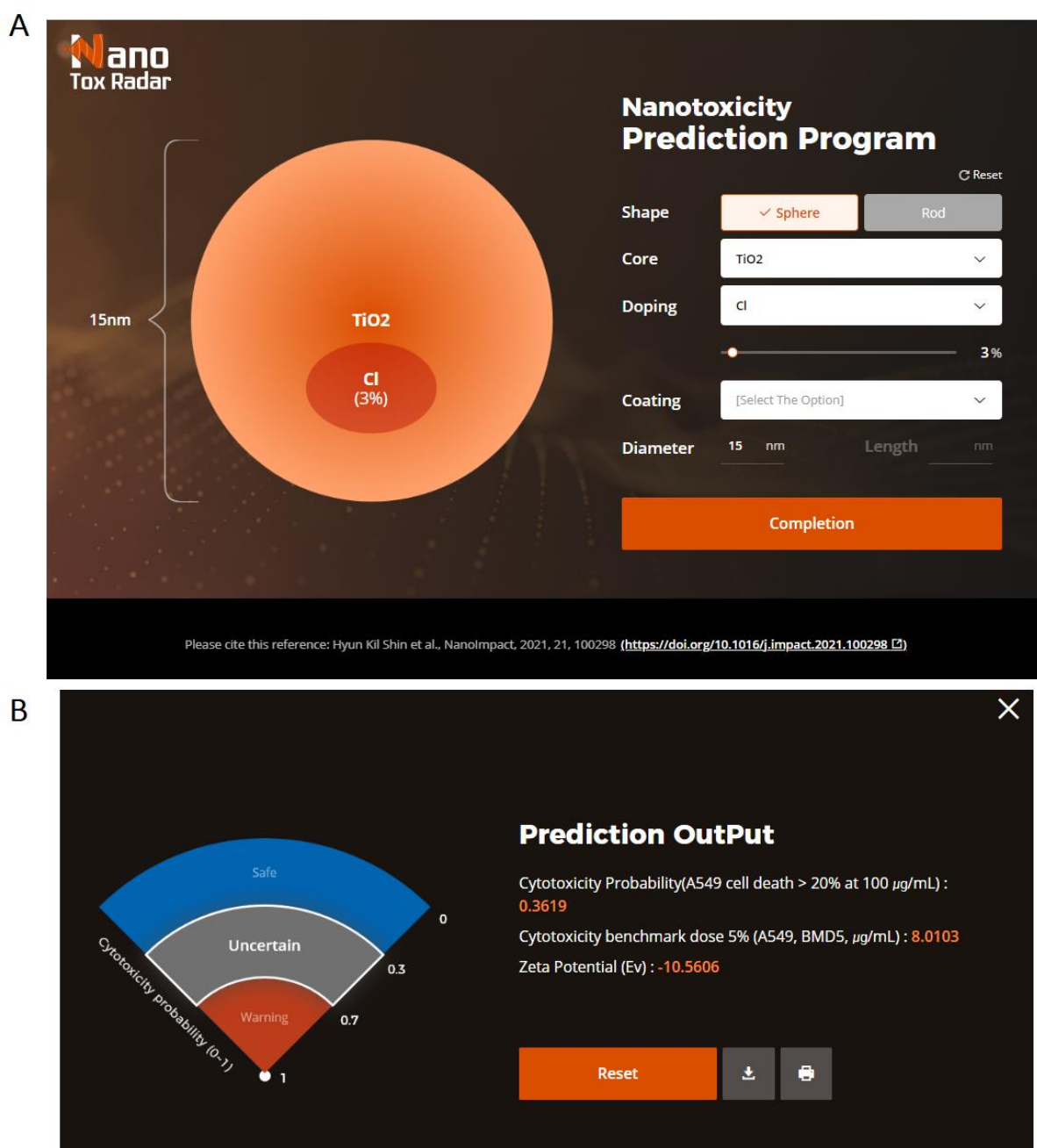


Figure 48: The NanoToxRadar tool web interface. [A] The users can tune the input parameters for the NMs of interest including their shape, their core, doping and coating parameters, and their size. [B] Upon execution the application presents the cytotoxicity probability, the cytotoxicity BMD5 and the zeta potential values of the input NM, along with an indication of their safety based on the cytotoxicity probability values.

The NanoToxRadar was employed for spherical and rod TiO<sub>2</sub>-based NMs to assess their cytotoxicity and zeta potential as functions of their diameter and their doping material. Unfortunately, the application does not quantify the reliability of the predictions based on the DoA of the background models. Thus, we tried to remain in the model domain based on the training TiO<sub>2</sub> NMs (supplementary data file from Shin *et al.*<sup>201</sup>). The 15 data points were for spherical TiO<sub>2</sub> NMs with diameter 6, 7, 10, 15 and 20nm,

without any coating material, undoped or doped with Cl (3%) or Na (3%). The results are presented in Figure 49 and Figure 50.

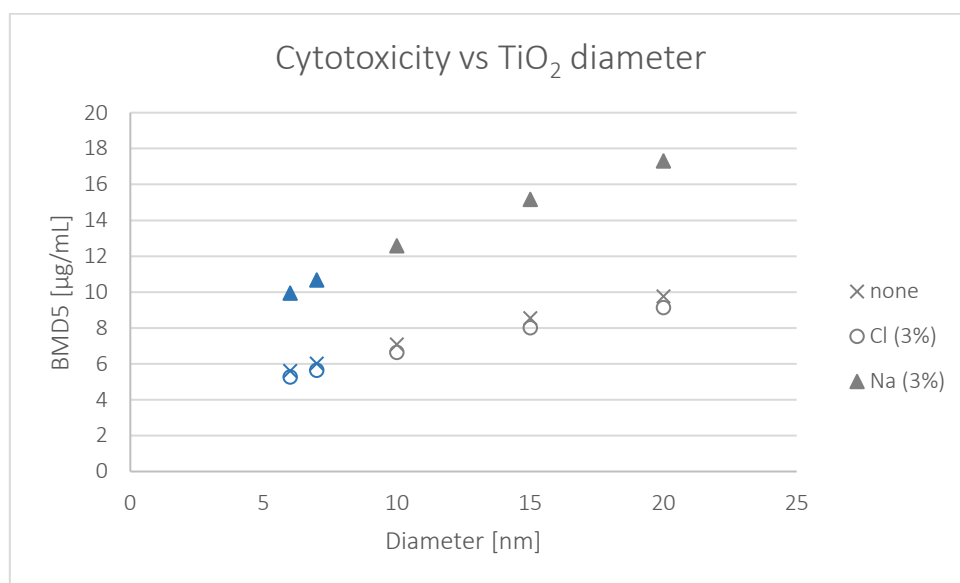


Figure 49: Cytotoxicity predicted values for different TiO<sub>2</sub> NM, as a function of their diameter. The numerical values represent the predicted BMD5 value, whereas the colour of the markers corresponds to the predicted cytotoxicity probability class: the blue colour corresponds to safe NMs and the grey colour to uncertain ones. Different marker shapes represented different doping parameters.

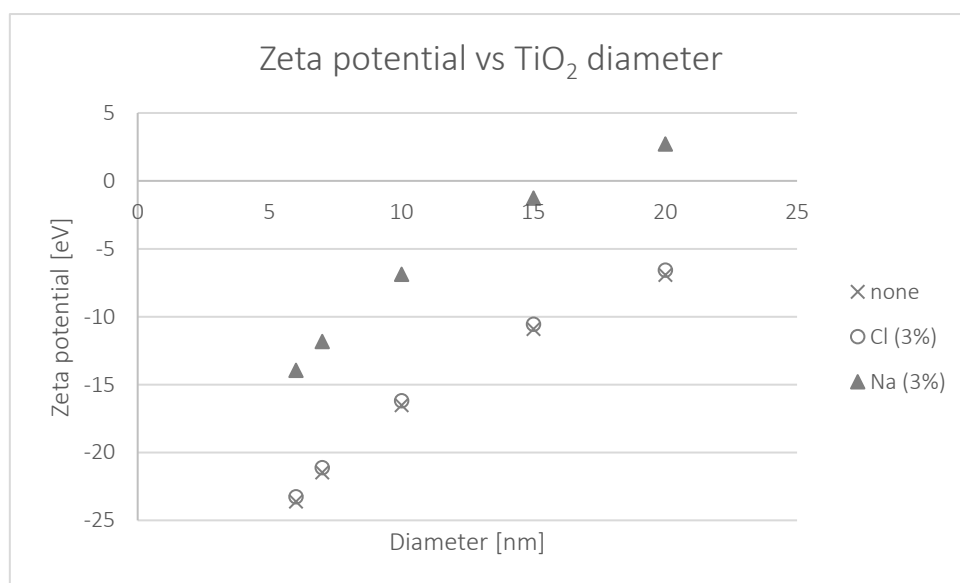


Figure 50: Zeta potential predicted values for different TiO<sub>2</sub> NM, as a function of their diameter. Different marker shapes represented different doping parameters.

#### 7.2.2.3.1 Discussion

NanoToxRadar is simple and straightforward to use to predict the cytotoxicity and the zeta potential of various NMs with different combinations of core, doping, and coating parameters. It can be used to predict the properties of NMs as an initial screening prior to their experimental evaluation or in an SbD study prior to their actual synthesis. A major flaw of the tool is the absence of reliability estimates for the predictions based on the applicability domain of the background models. According to the source publication<sup>201</sup>, the three models (two for the cytotoxicity predictions and one for the zeta potential prediction) were not trained on the same data, thus, different applicability domains exist for each

model. Additionally, polystyrene NMs were not included in the BMD5 training set but the model makes predictions for such NMs. Clearly, a reliability indication is needed for each predicted endpoint. Paradoxically, the applicability domain of each model is assessed in the publication but is not included in the tool. Moreover, the tool does not have a training material and it is not described well in the relevant publication of Shin *et al.*<sup>201</sup>. The stakeholders will require extra information on how to interpret results, and how the colour-coded cytotoxicity probability limits are derived<sup>i</sup> before using the tool. These improvements are required before its use in real-case scenarios and before it is trusted by stakeholders.

#### 7.2.2.4 INSIdE nano

The INSIdE nano web tool (<http://inano.biobyte.de/>) depicts and highlights connections between phenotypic entities, based on their effects on genes, for four categories of elements (NMs, drugs, chemicals and diseases)<sup>167</sup>. For each category, a list of associated genes is provided, derived from different sources. For NMs, gene expression data in different human cell types exposed to diverse NMs are available from the NanoMiner project. For drugs, gene expression data for drug treatments was retrieved from the Connectivity Map (<https://www.broadinstitute.org/connectivity-map-cmap>). Finally, for diseases and chemicals, gene interaction data was extracted from the Comparative Toxicogenomics Database (<http://ctdbase.org/>) and was manually curated.

The tool provides two ways to analyse the data set: visualisation of the phenotypic network; and visualisation of clustering of the phenotypic entities/categories of elements. For NMs, network construction is based on an ordered list of genes resulting from differential expression analysis. Pairwise similarity is then calculated between all entities and the respective values are used to build a weighted undirected network, where the nodes are the phenotypic entities and the similarity between them represents the weight of the edge (line). Positive and negative correlations between nodes-entities are encoded by the colour of the edges. Serra *et al.*<sup>167</sup> highlighted the remarkable association of metal and MeOx NMs and neurodegenerative disorders. The cluster analysis panel allows investigation of how the four entities are grouped and which genes are most important to each. In addition, query analysis (simple or conditional) is possible through the tool. This functionality allows the users to examine connections of a specific element (e.g., a NM), highlight its position in the ranking of neighbours, it indicates whether the connection is known in the literature, and specifies the connection distributions. Here, the different functionalities of the INSIdE nano tool are demonstrated for TiO<sub>2</sub> NMs.

##### 7.2.2.4.1 Browse

The user can view and interact with the network using the network browser tool. Our analysis began by using term TiO<sub>2</sub> in the search field of the network browser, then selecting for visualisation the top 5% of the interactions with diseases. The tool produces a network of the interactions between the TiO<sub>2</sub> NMs and the 7 diseases included in the top 5% of interactions. Positive correlations/interactions between nodes (TiO<sub>2</sub> NMs and diseases) are highlighted in red colour, whereas negative interactions in green colour. The weight of the edges between nodes indicates the degree of correlation between the nodes. By studying the results, we can derive meaningful conclusions for the query interactions e.g., that the TiO<sub>2</sub> NMs have negative interactions (green colour) on capillary leak syndrome (a cardiovascular disorder) and positive interactions (red colour) with the Tourette syndrome (a neuropsychological disorder) (Figure 51).

---

<sup>i</sup> The respective model as presented in the publication of Shin *et al.*<sup>201</sup> is trained on a dataset with binary labels for the cytotoxicity class.

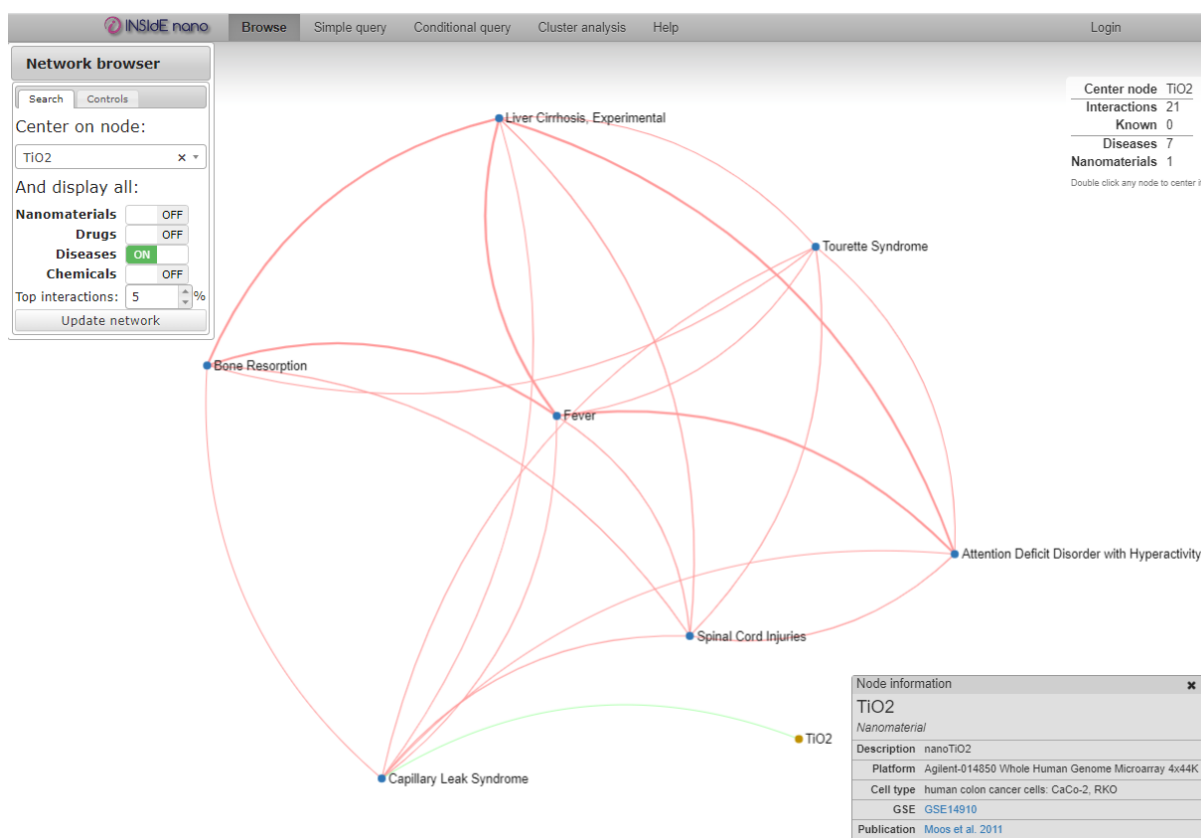


Figure 51: Results of the network browser for the top 5% interactions of TiO<sub>2</sub> NMs with diseases. Positive and negative interactions are colour coded (red and green respectively). By clicking on the different nodes, a pop-up window is presented with information about the node.

#### 7.2.2.4.2 Simple query

To further analyse the top 5% interactions of the TiO<sub>2</sub> NMs with the four categories of elements, a simple query was performed. The above parameters were inserted in the respective field and the resulting connections were displayed separately for the different groups of elements (see Figure 52). Again, the positive and negative interactions are colour-coded and, by clicking on the different elements, an information pop-up window is presented with further information and links. If the interactions-connections are known in literature, this is highlighted in a separate column. Additional data on the connection distribution and weights are presented in the next tabs (Figure 53 and Figure 54). However, if we insert the same parameters used in the “Browse” functionality above, we observe that TiO<sub>2</sub> NMs have positive interactions (red colour) with both the Tourette’s and the capillary leak syndrome (Figure 52). This inconsistency must be investigated further.

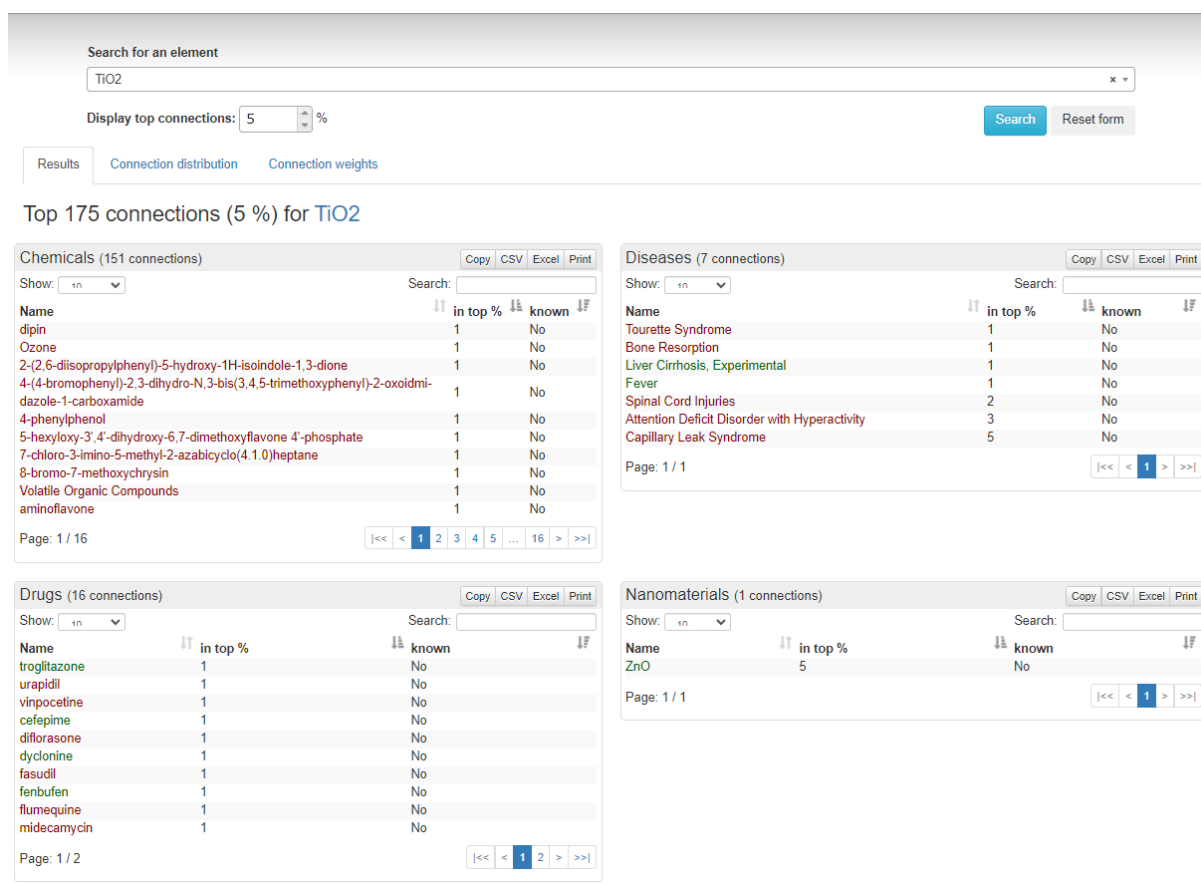


Figure 52: Simple query results for the top 5% interactions of TiO<sub>2</sub> NMs with chemicals, drugs, diseases and other NMs. Positive and negative interactions are colour coded (red and green respectively). By clicking on the different elements, a pop-up window is presented with information about the element. The results can be printed or exported in spreadsheet file format.

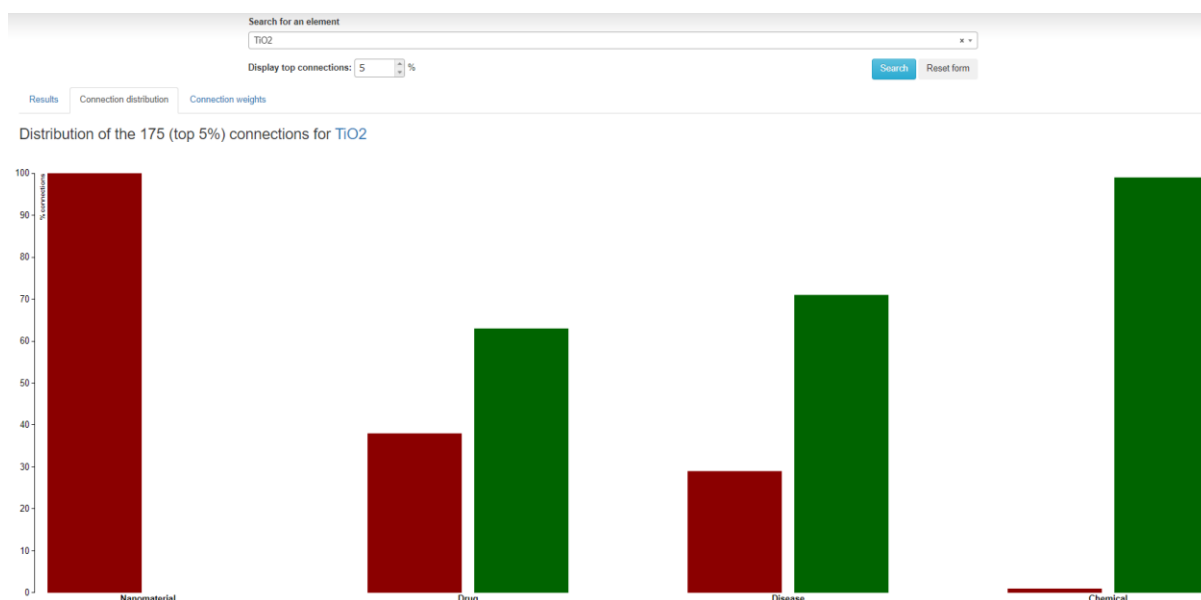


Figure 53: Simple query results for the top 5% interactions of TiO<sub>2</sub> NMs with chemicals, drugs, diseases and other NMs. The distribution of the positive (red) and negative (green) interactions considering the elements of each category is presented in graphical format.

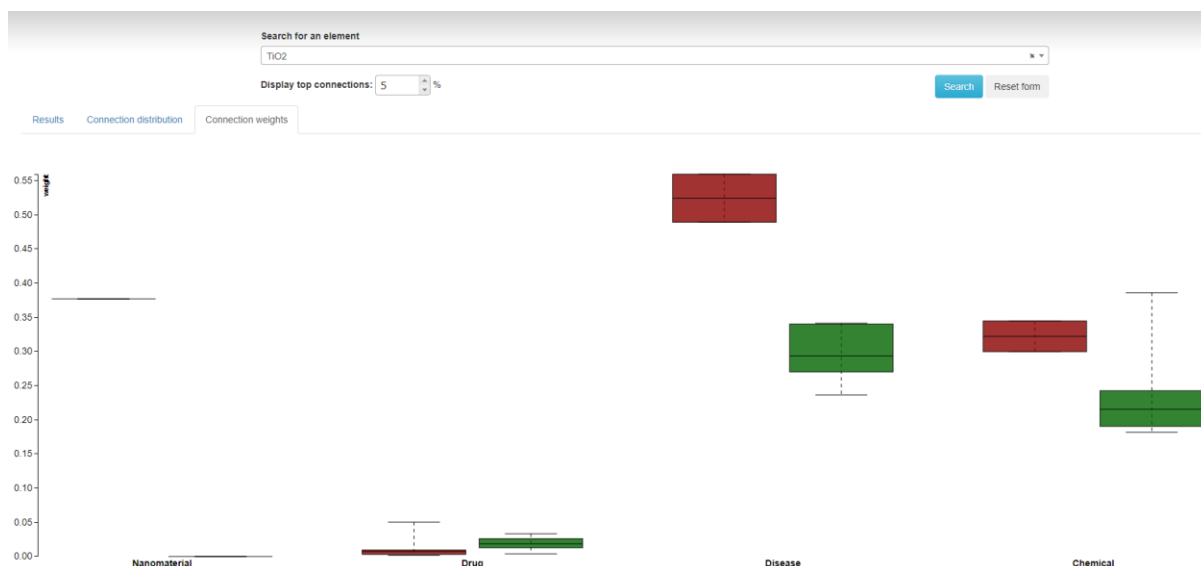


Figure 54: Simple query results for the top 5% interactions of TiO<sub>2</sub> NMs with chemicals, drugs, diseases and other NMs. The weights distribution for each category is presented in boxplots.

#### 7.2.2.4.3 Conditional query

A conditional query analysis was performed next that included TiO<sub>2</sub> NMs, all drugs in the training datasets, and Tourette's and capillary leak syndrome (Figure 55). Again, the top 5% interactions between elements were selected, and the minimum number of connected elements (neighbours) and of the elements in the cliques was set to 2. The cliques in graph theory are sub-graphs of the original network where all the nodes are connected to each other. Here, there are structures of three or four heterogeneous nodes (a disease, a drug, a chemical and an ENM) completely interconnected by strong patterns of similarity or anti-similarity<sup>167</sup>.

The figure shows the 'Conditional query' interface of the INSIDE nano tool. It includes sections for 'Element selection' and 'Query options'. In 'Element selection', 'Nanomaterials' is set to OFF, 'Drugs' is set to ON, 'Diseases' is set to OFF, and 'Chemicals' is set to OFF. The 'Individual nodes' section shows 'TiO2' selected under Nanomaterials, and 'Tourette Syndrome' and 'Capillary Leak Syndrome' selected under Diseases. In 'Query options', 'Top interactions' is set to 5%, 'Minimum connected elements' is set to 2, and 'Minimum elements in cliques' is set to 2. There are buttons for 'Run conditional query' and 'Reset form'.

Figure 55: Input parameters for the conditional query analysis.

The generated results by the INSIDE nano tool are presented in Figure 56 to Figure 58. The clique type list is presented first, based on the categories of elements that they contain. By selecting one of the cliques, additional information is presented such as pathway enrichment analysis results and gene enrichment analysis (Figure 56). The network can also be displayed (Figure 57) and users can zoom in to specific regions to inspect results, as the network may comprise a large number of branches. Finally, the stakeholders can visualise the number of connections between the elements in the different

**Query** Results Statistics

Your query matched **58** cliques, with a total of **62** nodes and **174** interactions (**0** known) in the network.

---

### Clique types:

Drug - Disease - Chemical (56)
Nanomaterial - Disease - Chemical (2)

### Matching cliques:

Filter by known interactions: at least one interaction is known: ☐ OFF

Filter by correlation: Nanomaterial ☒ is positively correlated with Nanomaterial ☐ Filter

Show:

Nanomaterial	Disease	Chemical	Known interactions
TIO2	Tourette Syndrome	N,N-bis(2-hydroxybenzyl)ethylenediamine-N,N'-diacetic acid	0
TIO2	Tourette Syndrome	alsikiren	0

Page: 1 / 1 (2 cliques)

### Clique information:

**Nanomaterial:** TIO2

**Disease:** Tourette Syndrome

**Chemical:** hydroxybenzyl ethylenediamine-N,N'-diacetic acid

**Pathway enrichment analysis**

There are 12 genes in 2 entries which are affected by one of the clique members.

Jaccard index threshold: 0.45 Update

Show: all

Category	Entry	Genes
BioCarta	BIOCARTA_TPO_PATHWAY	1
Reactome	REACTOME_SHC_RELATED_EVENTS	11

Page: 1 / 1

Highlight clique in the network

Gene enrichment analysis

Pathway enrichment analysis

Pubmed search

Query Results Network Statistics

**Network control**

Zoom

Network layout

Force atlas: [Start](#)

Force link: [Run](#)

Fruchterman-Reingold: [Run](#)

Display

Weighted nodes ☒

Weighted edges ☒

Filter

Interactions

positive ☒

negative ☒

only known ☐ OFF

Weight > 0

Nodes

Nanomaterials ☒

Drugs ☒

Diseases ☒

Chemicals ☒

[Reset all filters](#)



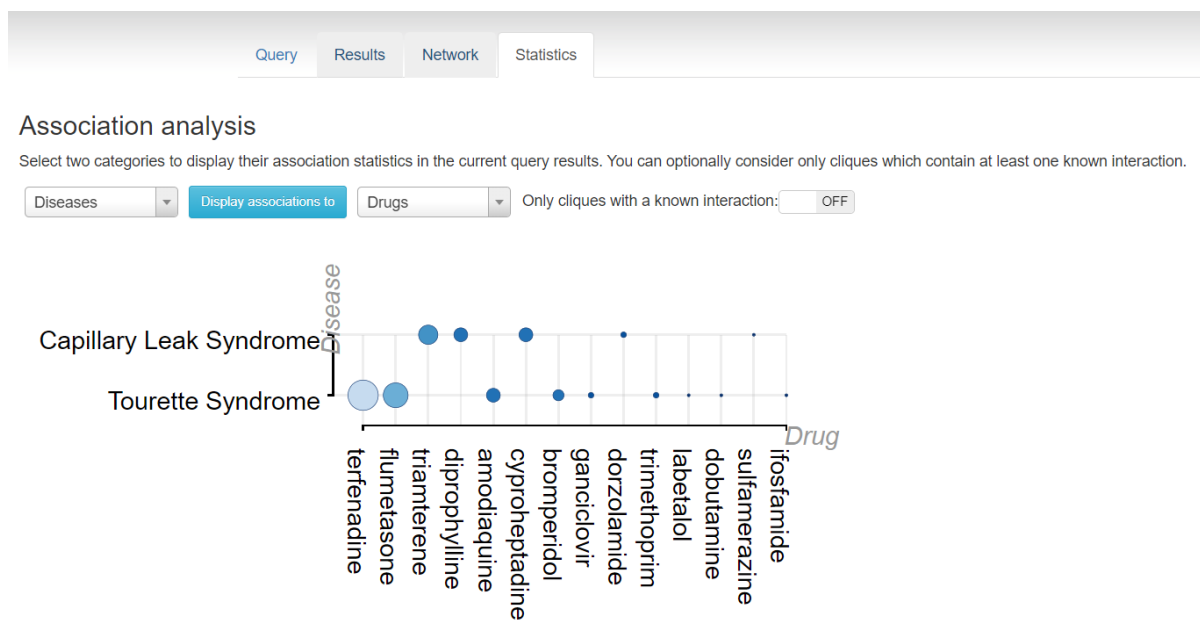


Figure 58: Results of the conditional query analysis (statistics) using the input parameters of Figure 55.

#### 7.2.2.4.4 Cluster analysis

The tool also provides cluster analysis to investigate different groups of NMs. TiO<sub>2</sub> NMs were chosen from the elements categories, and a list of up- and down-regulated genes from the differential expression analysis<sup>167</sup> are presented (Figure 59). The Gene Card webpage for each gene can be accessed by clicking on its name.

#### 7.2.2.4.5 Discussion

The INSIDE nano app offers many ways to explore a network of interactions between NMs, drugs, chemicals, and diseases. This network was built based on the systematic analysis of transcriptional mechanisms-of-action for these four entities. Due to its complexity, the use of the tool is not straightforward (like other tools in this review) and a thorough study of the instructions is required to understand how to exploit the different functionalities. The tool is supplemented by a comprehensive tutorial/help page that can guide stakeholders. However, some terms used, such as “cliques”, should be described in simpler language. In addition, the publication states that positive and negative MoA similarities in the interaction network (e.g., Figure 51) imply that genes are altered in the same or opposite direction, respectively, by both the phenotypic perturbations. Nonetheless, the lack of more explanation leaves room for “intuitive assumptions” that may be false, such as the belief that “positive” interactions between elements are “beneficial” and “negative” interactions are “detrimental”.

This application permits stakeholders not familiar with omics data analysis and integration to obtain results of different studies (e.g., gene expression data from cells exposed to NMs and from drug treatments, chemical-gene and disease-gene interactions data) and use information encoded in the data without any pre-requisites. The tool can be also used in a read-across framework because the networks generated are based on pairwise similarities between the different phenotypic entities under the same property space. Nonetheless, the challenge for this application is to ensure that it won't be outdated. As new data on the NMs mode-of-action are produced, it should be analysed and integrated into the tool to guarantee utility over time. A revision of the tool may be also needed to track down and correct the errors and inconsistencies (such as that mentioned earlier).

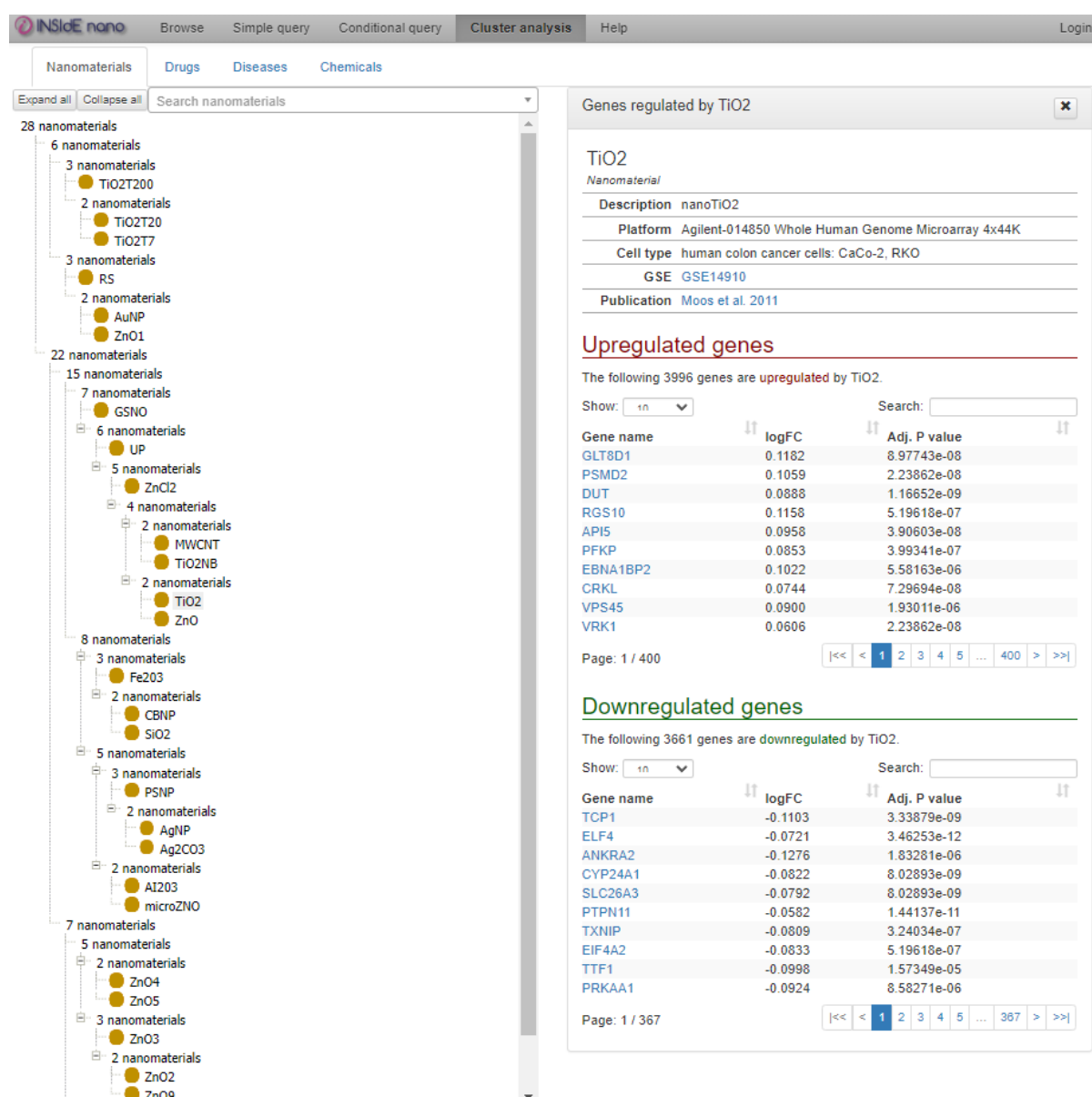


Figure 59: Cluster analysis results for the NMs category of elements.

### 7.2.2.5 Tool for assessment of human exposure to nanomaterials (multi-box aerosol model)

The multi-box aerosol model (<https://aerosol.cloud.nanosolveit.eu/>) is an easy-to-use web application for assessment of human exposure to NPs in indoor environments. It estimates the NP concentration in an indoor environment, e.g., a workplace laboratory or office, by dividing the space into smaller areas (near and far from the NPs emission source)<sup>161</sup>. The prediction depends on the geometrical characteristics of the room; therefore, the computational domain must be defined by the user (geometrical layout of the room, Near Field (NF) and Far Field(s) (FF), cuboid or cylindrical). A sketch of the computational domain is provided, based on the user entries. After that, in the scenario description section, the user needs to define the NPs emission rate either by choosing a specific scenario or by providing a distributed emission rate. Then, simulation inputs such as the modelling time, the time the source is active, and the number of repetitions must be specified, and a change of physical parameters (density, pressure, temperature, friction velocity) can be given. The output of the simulation consists of the evolution of the NP concentration in a particular or in all areas for different NP sizes (bins).

The multi-box aerosol tool was not demonstrated stand-alone but was integrated into a broader risk assessment scenario. Specifically, we predicted the biodistribution of TiO<sub>2</sub> NMs in the human respiratory system and other regions of the human body following exposure to NMs in indoor environments as part of an integrated approach to testing and assessment (IATA)<sup>j</sup>. This was achieved with the combination of three different tools that address different aspects of this scenario. The multi-box aerosol model is used to predict indoor air concentrations of NMs, the lung-exposure model predicts the accumulated NM mass in the human respiratory system, and the exposure PBPK model predicts the NM accumulated mass in different body compartments (Figure 60).

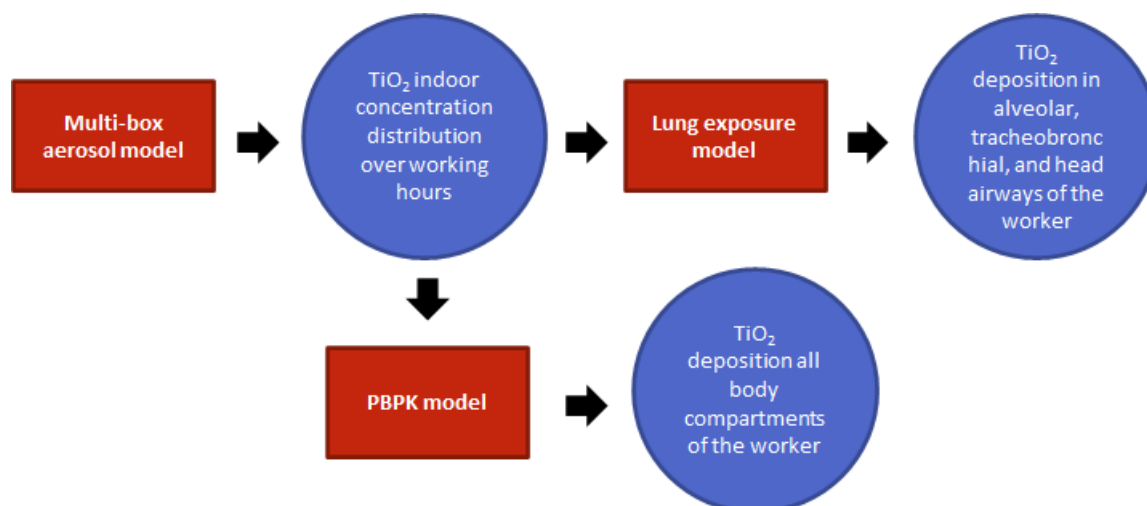


Figure 60: Schematic representation of the combination of different tools to simulate the biodistribution of TiO<sub>2</sub> NMs in the human respiratory system and other regions of the human body following exposure to NMs in indoor environments.

To predict the airborne concentration of the TiO<sub>2</sub> NMs in an indoor environment, the computational domain is a closed room where the NMs source is located in the NF area and the FF area is considered the exhaust area of the system. The room size is 4 × 4 × 4 m<sup>3</sup> and the NF size is 1 × 1 × 1 m<sup>3</sup>. The flow rate between NF and FF is set to 10 m<sup>3</sup>/min and the exhaust flow rate is set to 5 m<sup>3</sup>/min. From the available NMs, TiO<sub>2</sub> with a geometric mean diameter of 22nm and density of 4230kg/cm<sup>3</sup> was selected. A 6-hour work shift (1800s × 12 repetitions) was selected, because larger modelling times lead to server timeout. The NMs source was selected to be active for 15 minutes (900s) every time. The input parameters and the generated airborne distribution are presented in Figure 61. After the execution of the tool, the results were downloaded for use in the lung exposure model that predicts the deposited dose of the NMs in various parts of the human respiratory system, and in the PBPK model that predicts NM concentrations throughout the body.

<sup>j</sup> Integrated approaches to testing and assessment (IATA) provide a framework for combining information from different sources (experimental, *in silico*) for hazard characterisation of chemicals, including NMs, based on a weight of evidence approach.

User Guide

Computational domain

Number of FF(s): 1

Area	Schema	Starting point (from lower left corner)					Schema dimensions (m)				Flow to next area (m3/min)	Flow from inlet (m3/min)	Flow to exhaust (m3/min)			
Room	Cuboid	x	0	y	0	z	0	Length	4	Width	4	Height	4	0	0	0
NF	Cuboid	x	1.5	y	1.5	z	1.5	Length	1	Width	1	Height	1	10	0	0
FF1	Cuboid	x	0	y	0	z	0	Length	4	Width	4	Height	4	10	0	5

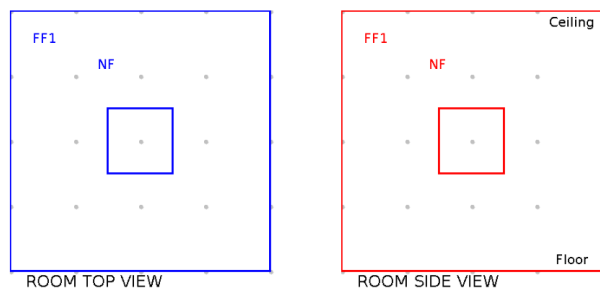


Figure 61: Input parameters of the multi-box aerosol model.

#### 7.2.2.5.1 Discussion

The prediction of the distribution of NMs in indoor environments through a user-friendly interface and its level of integration in other tools makes the multi-box aerosol model a valuable aid in risk and life cycle impact assessment of health effects of NMs. However, the model is slow to execute (e.g., for results presented here with 6 modelling hours, approximately ten computational/running hours were required<sup>k</sup>) and a major problem arose when longer simulation times were requested. Server timeouts and process interruptions occurred when longer modelling times were set as input, leading in loss of results and necessitating starting the computations again. Next versions of the application should focus on reducing computational times and on other functionalities useful for stakeholders. For example, a bar visualising the progression of the calculations would permit the users to understand the duration of the process. Another feature that allowed pausing or checkpointing of calculations, would be very useful, allowing restarting of stalled calculations.

#### 7.2.2.6 Lung exposure dose calculator

The lung exposure dose calculator (<https://lungexposure.cloud.nanosolveit.eu/>) uses a biokinetics model to estimate short-term regional lung-deposited dose of inhaled inorganic engineered NPs in humans following acute exposure (e.g., the result of an accident)<sup>161</sup>. The first step is to provide data for the concentration of NPs over time using three alternative models for NP diameter distribution (theoretical, template, or multi-box aerosol). Next, two models can be used for the calculation of NP deposition in the human respiratory system: the international commission on radiological protection model (ICRP); and an advanced ICRP model. Users must provide detailed parameter values for the exposed person when the advanced ICRP model is selected. The web-service then calculates the acute doses (in mg) for three regions of the respiratory system (alveolar, tracheobronchial, and head airways).

Here, the results of the multi-box aerosol calculations were used to describe the concentration of NMs over time. Then the simple ICRP model was employed to compute the NM deposited dose in the human respiratory system after 4 and 6 hours of exposure (14400s and 21600s respectively) in the NF and the FF (Figure 62). The results for different workers and activity levels are presented in

<sup>k</sup> The web application was accessed through a personal computer (processor: AMD Ryzen 7 5800H with Radeon Graphics 3.20 GHz, RAM: 16 GB). The web application calculations are performed on a server.

Table 16 and

Table 17 and depicted schematically in Figure 63 and Figure 64. Clearly, most of the NM mass is deposited in the head airways (nose, nasal cavity, and pharynx) and is lower for sedentary exposure than during heavy exercise.

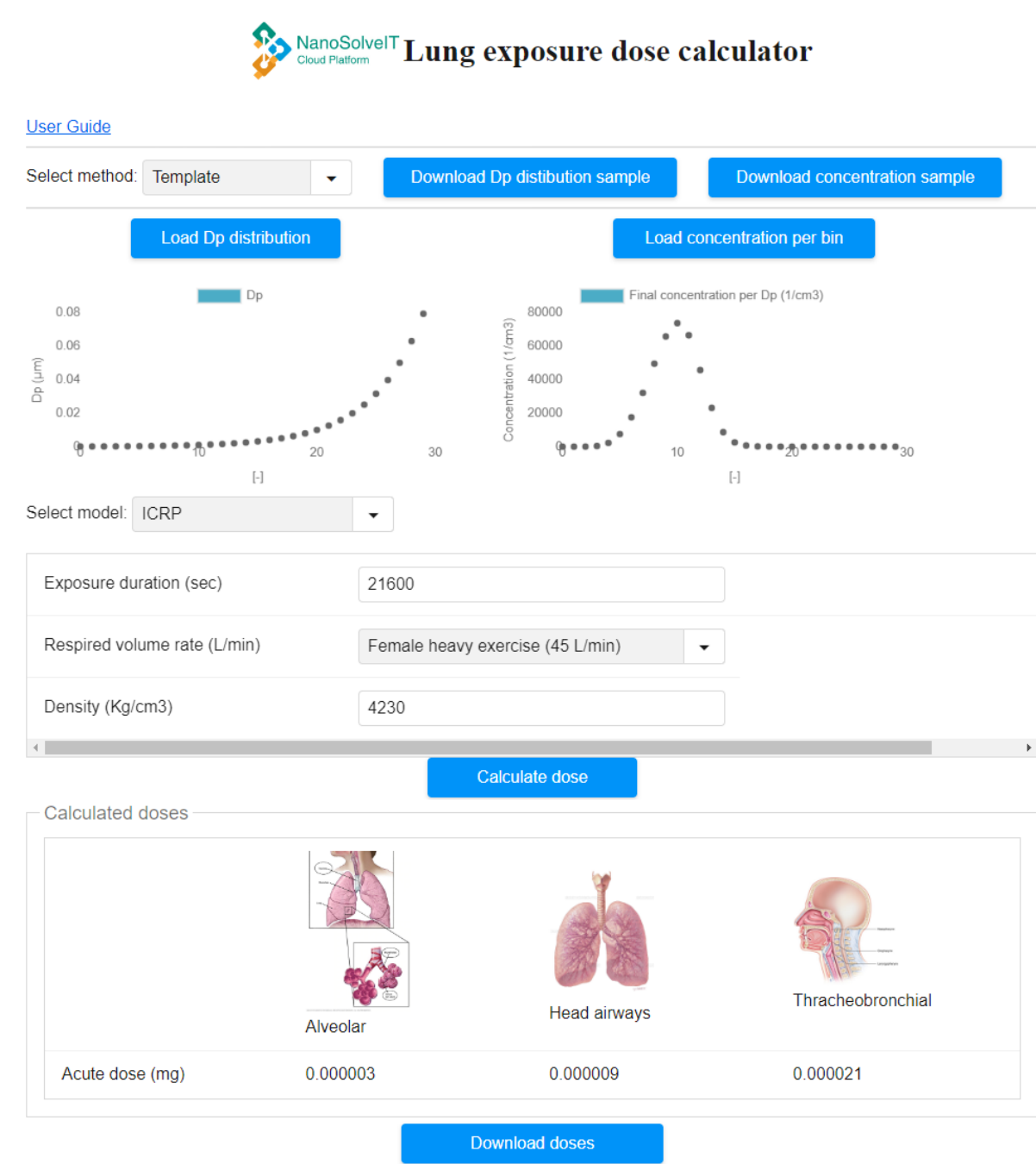


Figure 62: Deposited TiO<sub>2</sub> NMs (22nm) in the human respiratory system as produced by the lung exposure dose calculator using data for the airborne concentration of the NMs in the indoor environment from the multi-box aerosol model.

Table 16: Deposited dose in the respiratory system in the NF using the simple ICRP model.

Worker	Respired volume rate (L/min)	Exposure time [s]	Acute dose (mg)		
			Alveolar	Tracheobronchial	Head airways
Female	6.50 (Sitting)	14400	2.39E-7	8.16E-7	2.02E-6
Female	20.83 (Light exercise)	14400	7.67E-7	2.62E-6	6.46E-6
Female	45.00 (Heavy exercise)	14400	1.66E-6	5.65E-6	1.40E-5
Male	9.00 (Sitting)	14400	3.31E-7	1.13E-6	2.79E-6
Male	25.00 (Light exercise)	14400	9.21E-7	3.14E-6	7.75E-6
Male	50.00 (Heavy exercise)	14400	1.84E-6	6.28E-6	1.55E-5
Female	6.50 (Sitting)	21600	3.61E-7	1.23E-6	3.04E-6
Female	20.83 (Light exercise)	21600	1.16E-6	3.95E-6	9.75E-6
Female	45.00 (Heavy exercise)	21600	2.50E-6	8.53E-6	2.11E-5
Male	9.00 (Sitting)	21600	5.00E-7	1.71E-6	4.21E-6
Male	25.00 (Light exercise)	21600	1.39E-6	4.74E-6	1.17E-5
Male	50.00 (Heavy exercise)	21600	2.78E-6	9.48E-6	2.34E-5

Table 17: Deposited dose in the respiratory system in the FF using the simple ICRP model.

Worker	Respired volume rate (L/min)	Exposure time [s]	Acute dose (mg)		
			Alveolar	Tracheobronchial	Head airways
Female	6.50 (Sitting)	14400	2.37E-07	8.10E-07	2.00E-06
Female	20.83 (Light exercise)	14400	7.61E-07	2.60E-06	6.41E-06
Female	45.00 (Heavy exercise)	14400	1.64E-06	5.61E-06	1.38E-05
Male	9.00 (Sitting)	14400	3.29E-07	1.12E-06	2.77E-06
Male	25.00 (Light exercise)	14400	9.13E-07	3.12E-06	7.69E-06
Male	50.00 (Heavy exercise)	14400	1.83E-06	6.23E-06	1.54E-05
Female	6.50 (Sitting)	21600	3.58E-07	1.22E-06	3.02E-06
Female	20.83 (Light exercise)	21600	1.15E-06	3.92E-06	9.67E-06
Female	45.00 (Heavy exercise)	21600	2.48E-06	8.46E-06	2.09E-05
Male	9.00 (Sitting)	21600	9.61E-08	1.69E-06	4.18E-06
Male	25.00 (Light exercise)	21600	1.38E-06	4.70E-06	1.16E-05
Male	50.00 (Heavy exercise)	21600	2.76E-06	9.40E-06	2.32E-05

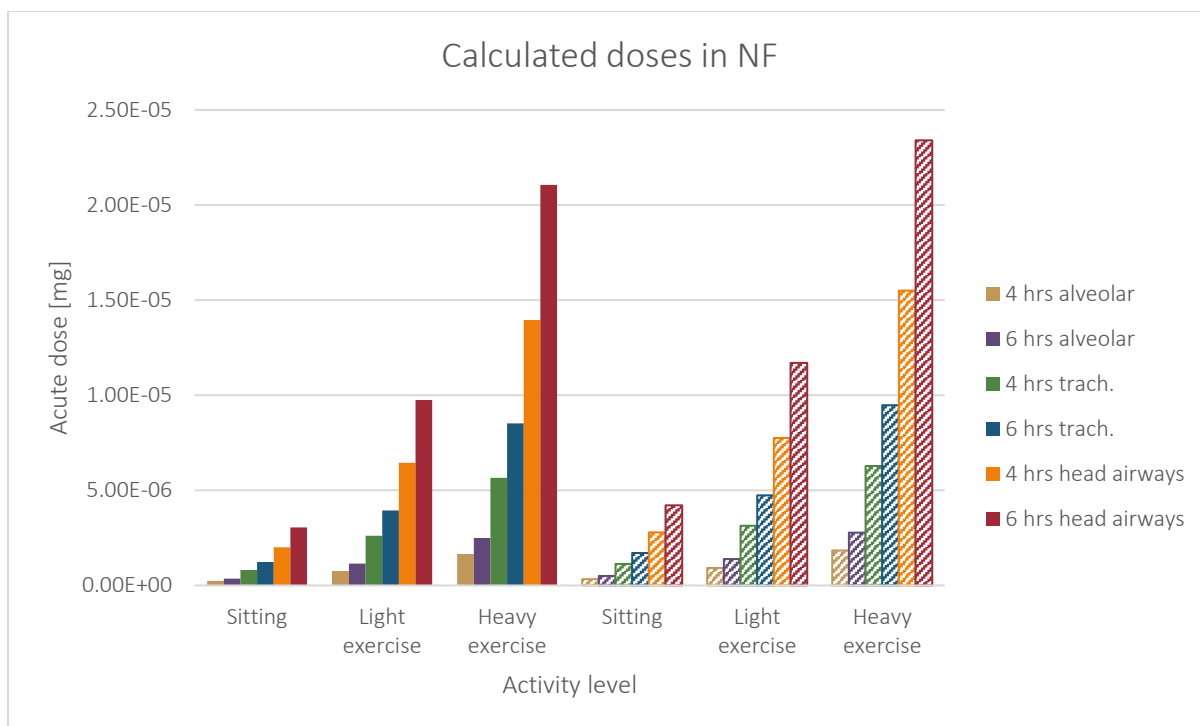


Figure 63: Calculated TiO<sub>2</sub> NM deposited doses in the different parts of the human respiratory system after 4 and 6 hours of exposure in NF. Solid bars refer to female whereas shaded bars refer to male workers.

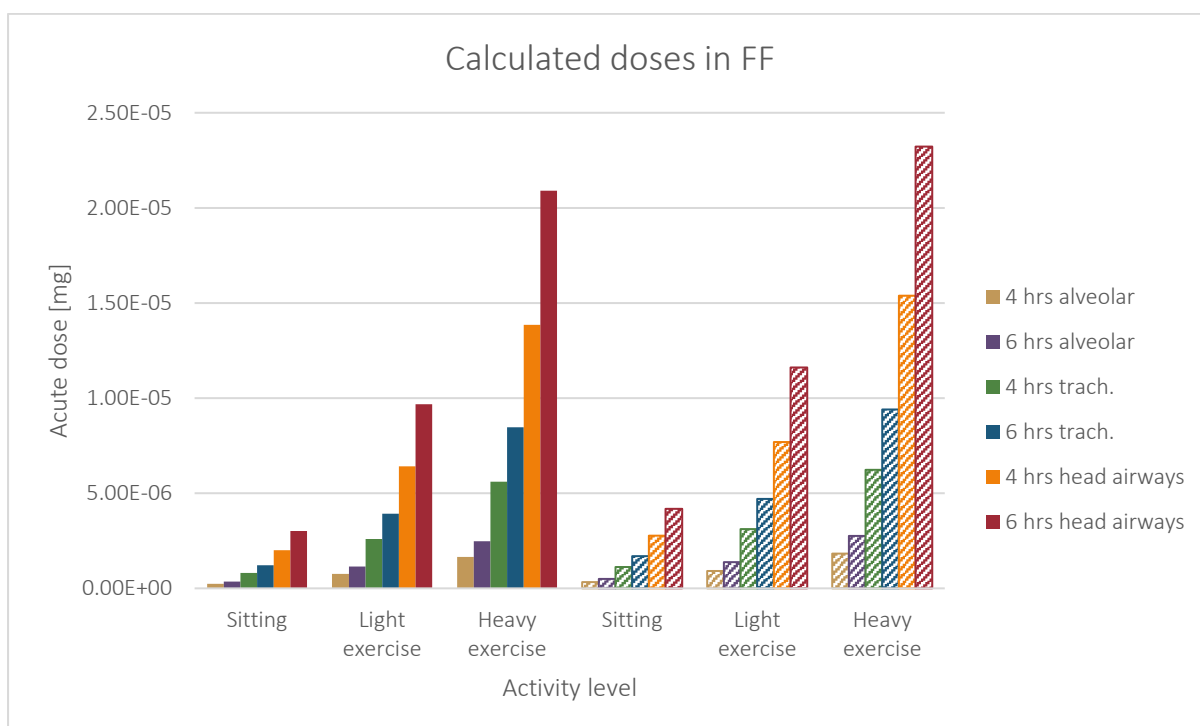


Figure 64: Calculated TiO<sub>2</sub> NM deposited doses in the different parts of the human respiratory system after 4 and 6 hours of exposure in FF. Solid bars refer to female whereas shaded bars refer to male workers.

#### 7.2.2.6.1 Discussion

The lung exposure dose calculator is simple to use and produces results that can be easily integrated in a risk assessment study for the effects of human exposure to NMs. A minor issue is that the tutorial is rather technical. To make the tool more useful to stakeholders, the tutorial should be improved to include more details on the input parameters in more simple language.

### 7.2.2.7 PBPK models and integration with the occupational exposure model

This web application uses PBPK modelling to simulate the NP biodistribution in humans due to inhalation exposure (<https://exposurepbpk.cloud.nanosolveit.eu/>)<sup>161</sup>. The application consists of an external and an internal model for inhalation route exposure to TiO<sub>2</sub> NMs. External exposure scenarios can be defined directly by the user by uploading a custom-made exposure scenario, or it can be simulated through the multi-box aerosol model. For the internal exposure module, the user provides information on the weight of the worker and the exposure duration (the last point of the simulation time vector in hours). The mass-time profiles are generated by the PBPK model for various body compartments (heart, spleen, brain, kidneys, alveolar, upper, and lower respiratory, liver, tracheobronchial, blood, skin, lung interstitium and capillaries and, rest of the body). The user can select from the drop-down menu the desired compartment and observe the NP deposition in this organ/tissue as a function of time. As the background PBPK model incorporates the effects of clearance processes (removal of the NMs from the human body), this tool extends the lung exposure dose calculator and produces more realistic biodistribution simulation results, especially for longer-term exposures.

To start the analysis, the results of the multi-box aerosol calculations in NF were uploaded to describe the external exposure scenario. Then the weight of the worker and the last point of the simulation time vector (6 hours) were adjusted as needed for the PBPK model. Upon execution, the mass-time profile for each organ was produced as presented in Figure 65. The deposited dose at the last point of the simulation for different worker weights is presented schematically in Figure 66. These show that higher deposited doses occur in the respiratory system.

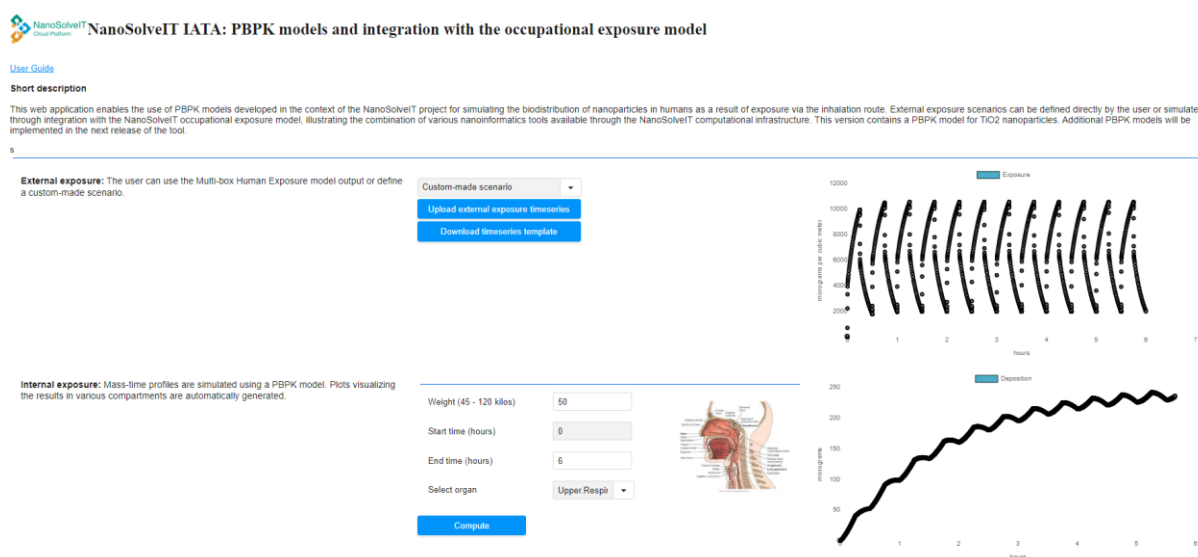


Figure 65: TiO<sub>2</sub> NMs (22nm) mass-time profile in the different organs/compartments of the human body (here in the upper respiratory system), as produced by the exposure-PBPK model.



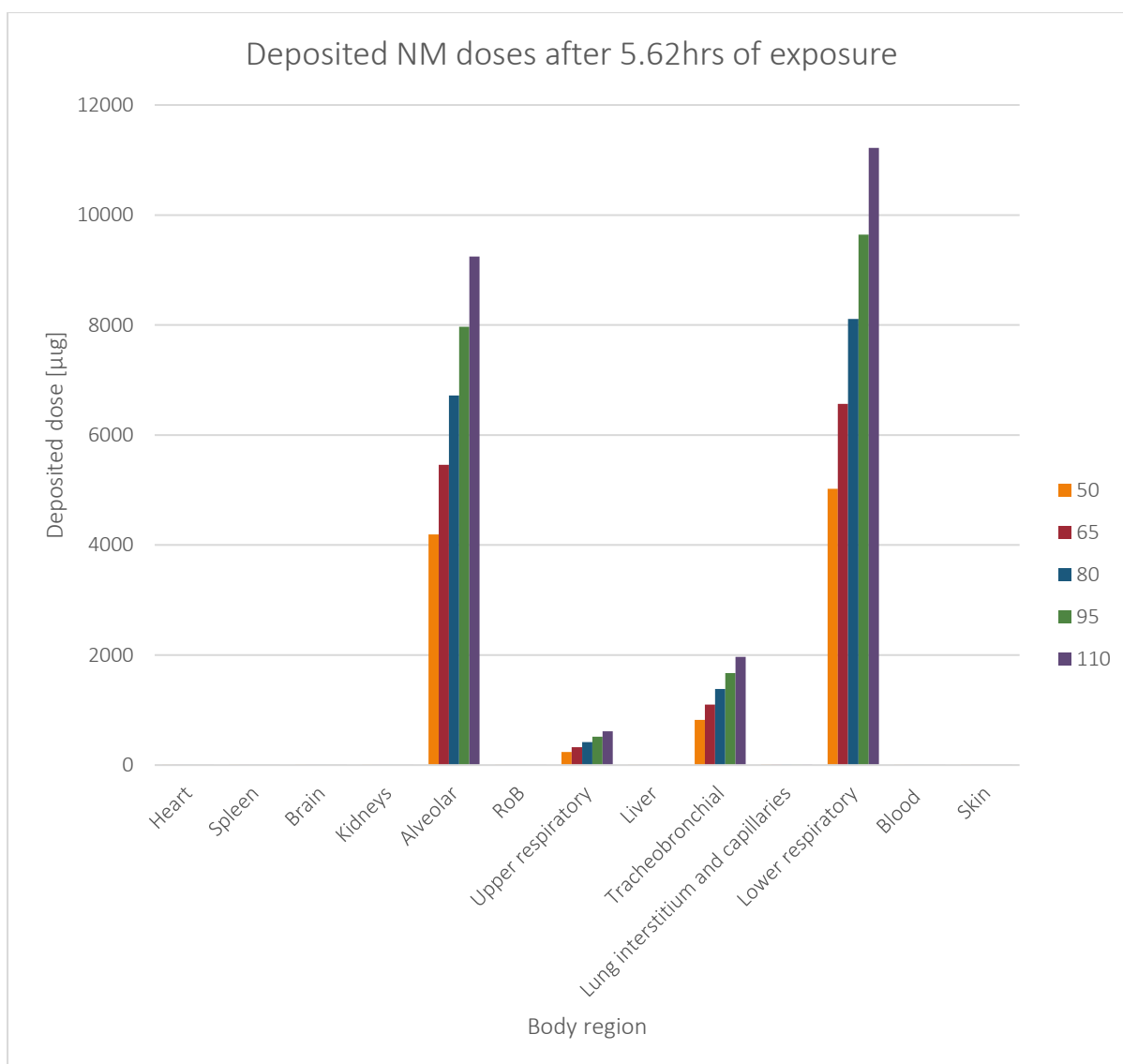


Figure 66: Deposited TiO<sub>2</sub> NMs (22nm) doses in different body regions/organs for different body weights (in kg) after 5.62hrs of exposure to the NM source according to the multi-box aerosol model.

#### 7.2.2.7.1 Discussion

The tool is simple to use, it exploits the results of the multi-box aerosol model, or it can be used with custom-made exposure scenarios to derive mass-time profiles of the deposited NMs in different organs/regions of the human body. These functions allow it to be used in combination with other risk assessment studies for regulatory purposes. Nevertheless, revisions are needed to allow full exploitation of the tool and to improve the user experience. For example, the mass-time profiles cannot be downloaded and stored in the user's computer. Thus, users cannot post-process their results (e.g., inspect the dose in intermediate time steps, present the mass-time profiles on different organs in the same plot, etc.). In addition, the simulations must be repeated every time the stakeholders want to use/present the specific biodistributions. Another issue is that the simulations did not span the entire six hours (the last time point was at 5.62 hours). Even when specifying an "end time" of 8-hours, the simulation still stopped at 5.62 hours. This prevents study of longer-term NM biodistribution (e.g., for a day or a week-period). Another operational issue is that server timeouts occurred and interrupted the calculations. Finally, the PBPK model was developed using biodistribution data for TiO<sub>2</sub> NMs of 22nm size and, therefore, simulations for TiO<sub>2</sub> NMs with similar diameter are likely to be the most reliable<sup>161</sup>. The model should be extended to a more diverse set of NM types and sizes.

## 7.3 Case study 3: Combination of nanoinformatics models into SAPNets

### 7.3.1 Scope of the third case study

The SAPNets as described by Rybińska-Fryca *et al.*<sup>112</sup> are a “series of mutually dependent predictive models” where the prediction of an endpoint is based on descriptors predicted by other “meta-models”. The development of a SAPNet is carried out in three phases. A NM’s structural characterisation, physical and chemical characteristics, and toxicity are all compiled first. Next, by creating “meta-models”, the most pertinent intrinsic and extrinsic features that are likely to affect NM’s toxicity are found. Eventually, these “meta-models” that outline the relationships between the most important characteristics of NMs and their unfavourable effects are created. The network is constructed in this manner, layer by layer, starting from the endpoint (such as toxicity or other properties) and ending with descriptors that define the NM structure. As a result, SAPNets can be directly used by non-specialists as they can easily perceive the structural characteristics that need to be changed (such as NMs size, shape, aspect-ratio, and type of coating) to produce a safe NM with desired properties. An example of a SAPNet is already presented in the second case study (NanoMixHamster tool). The concept of SAPNets can be further applied on the combination of existing models for the prediction of NMs properties.

In this case study, a nanoQSPR (quantitative structure–property relationship) model that expresses the zeta potential in a potassium chloride solution, as a function of physicochemical (zeta potential in water) and periodic table-based (periodic number, PN) descriptors is combined with another nanoinformatics model which provides the first model with calculated values of the zeta potential in water. In other words, the physicochemical descriptor used for the derivation of the desired endpoint, is not taken from an experimental study, but rather estimated *in silico*.

The workflow for the SAPNet analysed in this work was originally proposed (but not implemented) by Rybińska-Fryca *et al.*<sup>112</sup>. Here, we put into application, and we evaluate the possibility of combining models existing in the literature in an effort to exploit all the already available models and predict more nanoinformatics-related endpoints. Rybińska-Fryca *et al.*<sup>112</sup> suggest, for the prediction of the zeta potential in aquatic solution, the use of a QSPR model developed by Toropov *et al.*<sup>217</sup> developed using the Monte Carlo approach or the use of a nanoQSPR model developed by Mikolajczyk *et al.*<sup>142</sup> that requires information for physicochemical and molecular descriptors for the zeta potential predictions. In this case, we employed a different and simpler read-across model for the zeta potential in water which is based on periodic table descriptors and the NMs core size (MS<sup>3</sup>bD model)<sup>1</sup>.

### 7.3.2 Results of the third case study

The described SAPNet is presented schematically in Figure 67.



Figure 67: Graphical representation of the zeta potential in KCl prediction according to the SAPNet scheme.

**Model A:** Papadimitrakaki *et al.*<sup>1</sup> created the MS<sup>3</sup>bD (MScubed: Molecular, Size and Surface based Safe by Design) model, which is a fully validated *k*NN predictive model according to the OECD principles<sup>218</sup>, that calculates the zeta potential values in water in pH=7 ( $\zeta_{H_2O}$ ). The dataset used for developing the model contains twenty physicochemical properties for 69 ENMs<sup>219</sup> and can be found in the NanoPharos database (<https://db.nanopharos.eu/Queries/Datasets.zul?datasetID=3>). Moreover, among the twenty descriptors in the original dataset, five variables were selected to be included for model development. The properties required as input for the model are the type of coating, the NM core size measured with transmission electron microscopy (TEM) or scanning TEM (STEM), the ionic radius of the metal in the core of the NM, the absolute electronegativity ( $\chi_{abs}$ ) and the sum of metal electronegativity divided by the number of oxygen atoms present in the MeOx ( $\Sigma\chi/n_O$ ). The output results include the zeta potential prediction, and an indication whether the predicted value falls within the model's applicability domain, characterizing the prediction as either reliable or unreliable. The model is available through the Enalos Cloud Platform as a web service, which can be found at <https://mszeta.cloud.nanosolveit.eu/>, however training material is not available for this tool. The obtained values for  $\zeta_{H_2O}$  included in this study were derived online via this web service.

**Model B:** The obtained values of  $\zeta_{H_2O}$  for different NMs are used as input in a model developed by Wyrzykowska *et al.*<sup>98</sup> that predicts zeta potential in KCl solution ( $\zeta_{KCl}$ ) using the MLR. The model is developed with two standardised descriptors (using gaussian normalisation), zeta potential in distilled water ( $\zeta_{H_2O}$ , neutral pH of 7) and periodic number (PN), which indicates the number of electron shells in the MeOx NMs:

$$\zeta_{KCl} = 3.98 + 21.68 \cdot \zeta_{H_2O} + 7.88 \cdot PN$$

The original dataset based on which the linear model was built, was taken from Pathakoti *et al.*<sup>220</sup>. The two independent variables were selected using a genetic algorithm (GA-MLR), among a set of 62 physicochemical, quantum and periodic table-based descriptors, while the data were normalised and split into training and validation subsets, as presented in Table 18. To assess the quality of the model, authors followed the recommendations of the OECD<sup>119</sup> in model validation.

Table 18: Experimental data for model training and testing, including  $\zeta\text{KCl}$  (mV), and predicted by the nanoQSPR model<sup>98</sup> ( $\zeta\text{KCl}_{\text{predicted}}$  mV) for 15 MeOx NPs.

NP	Particle TEM size [nm]	$\zeta\text{H}_2\text{O}$ [mV]	PN	$\zeta\text{KCl}_{\text{exp}}$ [mV]	$\zeta\text{KCl}_{\text{predicted}}$ [mV]	Subset
$\text{In}_2\text{O}_3$	60	22.6	5	28.7	32	Training
$\text{NiO}$	14	26	4	26.8	26.5	Training
$\text{Al}_2\text{O}_3$	55	30.3	3	25.3	23.1	Training
$\text{La}_2\text{O}_3$	65	-3.6	6	22.3	12.8	Training
$\text{CuO}$	28	24.4	4	19.1	24.9	Training
$\text{ZrO}_2$	27	-6.9	5	4	1.9	Training
$\text{TiO}_2$	42	-10.7	4	-2.2	-9.4	Training
$\text{Bi}_2\text{O}_3$	144	-16.5	6	-4.9	0.2	Training
$\text{SnO}_2$	15	-21.1	5	-16.7	-12	Training
$\text{ZnO}$	71	-20.9	4	-24.9	-19.4	Training
$\text{SiO}_2$	20	-29.8	3	-33.7	-35.7	Training
$\text{CoO}$	55	17.5	4	26	18.2	Validation
$\text{Y}_2\text{O}_3$	38	16.3	5	17.9	24.6	Validation
$\text{Sb}_2\text{O}_3$	84	-20.7	5	-12.7	-11.6	Validation
$\text{V}_2\text{O}_3$	NA	-27.9	4	-32.6	-26.2	Validation

To begin with, we tried to reproduce the results of the Wyrzykowska *et al.* MLR model (also presented in Table 18) and, by applying the model on the data of  $\text{In}_2\text{O}_3$  the predicted  $\zeta\text{KCl}$  equals 533.35mV instead of 32mV. Same applies for all the predicted  $\zeta\text{KCl}$  values of the NMs in Table 18. In the publication of Wyrzykowska *et al.*<sup>98</sup> it is stated that the independent variables (descriptors) are auto-scaled, however no extra information is provided. For this reason, we reproduced the modelling steps in KNIME, including Gaussian normalisation of the descriptors, to ensure that we will use the correct model on the SAPNet development. In this course, regarding the applicability domain definition, a small issue arose for the leverage threshold ( $h^*$ ). For this parameter the authors considered to use the endpoint values to calculate the leverage ( $h^*$ ), except for the independent descriptors. This means that when an untested sample is examined, its leverage could not be calculated as the endpoint values are unknown. Thus, the correct leverage limit value ( $h^*$ ) is 0.55 instead of 0.82. In Figure 68: The William's plot for the graphical visualisation of the NMs relative position with respect to the leverage threshold ( $h^*$ ). Data are acquired from Table 18. No outliers exist in this case; thus, no extrapolated endpoint values are calculated for the validation set. Figure 68 a graph (William's plot) is presented that visualises the applicability domain of the model with respect to the  $h^*$  value, and thus helps to visualise interpolated and extrapolated predictions.

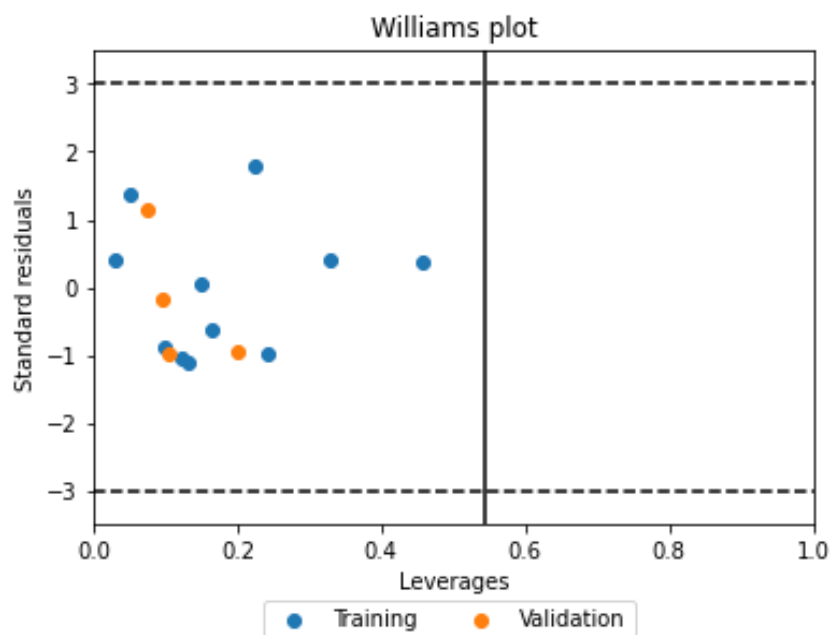


Figure 68: The William's plot for the graphical visualisation of the NMs relative position with respect to the leverage threshold ( $h^*$ ). Data are acquired from Table 18. No outliers exist in this case; thus, no extrapolated endpoint values are calculated for the validation set.

Wyrzykowska *et al.* used another dataset (available as supporting material in their publication) as an external set of datapoints for predictions of  $\zeta KCl$ . This dataset (

Table 19), comprises a list of experimentally derived values of the zeta potential in water for 44 different types of MeOx NPs, where the particle size is given and the  $\zeta$  values are unknown. Nonetheless, the conditions under which the zeta potential in water is measured are not included in the publication and we could not trace back the respective meta-data (e.g., pH of the solution) when searching on the data sources for the zeta potential values.

Table 19: External dataset with 44 experimental datapoints for zeta potential in water and the predictions of zeta potential in KCl according to Wyrzykowska *et al.*'s model<sup>98</sup>.

Row ID	MexOy NP	Hydrodyn. diameter [nm]	Particle size TEM [nm]	Engineered size [nm]	$\zeta_{H_2O,exp}$ [mV]	PN	$\zeta_{KCl,predicted}$ [mV]
1	ZnO	-	-	50	-55	4	-52.711
2	ZnO	-	-	50	-50	4	-47.8214
3	ZnO	-	-	60	-45	4	-42.9318
4	ZnO	-	-	1000	-44	4	-41.9539
5	ZnO	-	-	1200	-33	4	-31.1968
6	ZnO	-	-	1200	-32	4	-30.2188
7	ZnO	-	-	60	-30	4	-28.263
8	ZnO	-	-	70	-29	4	-27.2851
9	ZnO	-	-	50	-25	4	-23.3734
10	ZnO	-	-	1200	-31.8	4	-19.4617
11	ZnO	-	-	1000	-20	4	-18.4838
12	TiO <sub>2</sub>	-	-	30	-15	4	-13.5942
13	TiO <sub>2</sub>	-	-	30	-12	4	-10.6604
14	TiO <sub>2</sub>	-	-	45	-11	4	-9.6825
15	TiO <sub>2</sub>	-	-	125	-10	4	-8.7046
16	TiO <sub>2</sub>	-	-	125	-9	4	-7.7266
17	TiO <sub>2</sub>	-	-	125	-8	4	-6.7487
18	MnO <sub>2</sub>	-	-	-	-47.6	4	-45.4744
19	WO <sub>3</sub>	-	30,0-70,0	-	-9.1	6	7.3923
20	ZrO <sub>2</sub>	-	46.7	-	-8.5	5	0.3707
21	SnO <sub>2</sub>	-	46.1	-	-10.5	5	-1.5851
22	Sb <sub>2</sub> O <sub>3</sub>	-	90,0-210,0	-	-13.3	5	-4.3233
23	In <sub>2</sub> O <sub>3</sub>	-	29.8	-	-9.6	5	-0.705
24	La <sub>2</sub> O <sub>3</sub>	-	46	-	-12.8	6	3.774
25	V <sub>2</sub> O <sub>3</sub>	-	-	-	-22.8	4	-21.222
26	Al <sub>2</sub> O <sub>3</sub>	-	44	-	-20.2	3	36.4948
27	Fe <sub>2</sub> O <sub>3</sub>	-	32	-	-18.1	4	-16.6257
28	ZnO	-	71	-	-10.8	4	-9.4869
29	Y <sub>2</sub> O <sub>3</sub>	-	38	-	-10.7	5	-1.7807
30	TiO <sub>2</sub>	-	42.3	-	-9.6	4	-8.3134
31	SiO <sub>2</sub>	-	15	-	-8.1	3	-14.4549
32	CoO	-	<100	-	-3.4	4	-2.2503
33	Bi <sub>2</sub> O <sub>3</sub>	-	90	-	-2.3	6	14.0422
34	Al <sub>2</sub> O <sub>3</sub>	94.7	-	-	39.2	3	31.8008
35	Co <sub>3</sub> O <sub>4</sub>	99.2	-	-	23	4	23.5669
36	CuO	130	-	-	17	4	17.6993
37	Fe <sub>3</sub> O <sub>4</sub>	128	-	-	22.1	4	22.6867
38	MgO	1964	-	-	6.8	3	0.11615
39	Mn <sub>3</sub> O <sub>4</sub>	395	-	-	-14.3	4	-12.9096
40	Sb <sub>2</sub> O <sub>3</sub>	124.7	-	-	-24.2	5	-14.9827

<b>41</b>	SiO <sub>2</sub>	147.8	-	-	-33.5	3	-39.2941
<b>42</b>	ZnO	161.3	-	-	16.4	4	17.1126
<b>43</b>	TiO <sub>2</sub>	171	-	-	-13.4	4	-12.0295
<b>44</b>	WO <sub>3</sub>	62.8	-	-	-45.2	6	-27.9107

In order to incorporate the created SAPNet, a part of the datapoints from this dataset were used, first to calculate the zeta potential in water from the read-across model (model A) and compare the predictions to the respective experimental value, and then to predict the zeta potential in KCl from the nanoQSPR model (model B). The datapoints from



Table 19 that could be used in the SAPNet are the NMs in rows 19-24, 26-31 and 33, as they are the only ones with known core sizes measured with TEM (the necessary descriptor for model A). In case of the NMs in rows 19 and 22, the average diameter of the particles was used. Specifically, each uncoated NMs' core diameter is used as input in the MS<sup>3</sup>bD model, while the rest required data are periodic table descriptors. The  $\chi_{\text{abs}}$  were calculated by the formula provided in the respective publication<sup>11</sup>. It should be noted that solely  $\zeta_{\text{H}_2\text{O}}$  predictions characterised as “reliable” (interpolated predictions) were accepted and fed into model B, as the rest NM samples are located outside the applicability domain of the model A and predictions for them are considered less accurate (extrapolated predictions). Next, the predicted  $\zeta_{\text{H}_2\text{O}}$  values are used to predict the  $\zeta_{\text{KCl}}$  values, as summarised in Table 20.

Table 20: Zeta potential values in aquatic and KCl solutions for different types of NMs, derived following the SAPNets scheme.

NM type	Core diameter [nm]	$\zeta_{\text{H}_2\text{O}}$ , MS <sup>3</sup> bD model [mV]	Reliability of predicted $\zeta_{\text{H}_2\text{O}}$	$\zeta_{\text{H}_2\text{O}}$ exp. [mV]	$\zeta_{\text{KCl}}$ , SAPNet [mV]	Reliability of predicted $\zeta_{\text{KCl}}$
WO <sub>3</sub> , uncoated	50.0	-0.8	reliable	-9.1	15.5	reliable
ZrO <sub>2</sub> , uncoated	46.7	6.8	reliable	-8.5	15.3	reliable
SnO <sub>2</sub> , uncoated	46.1	-8.6	reliable	-10.5	0.2	reliable
Sb <sub>2</sub> O <sub>3</sub> , uncoated	150.0	-15.6	reliable	-13.3	-6.6	reliable
In <sub>2</sub> O <sub>3</sub> , uncoated	29.8	-11.0	reliable	-9.6	-2.1	reliable
La <sub>2</sub> O <sub>3</sub> , uncoated	46.0	-8.7	reliable	-12.8	7.8	reliable
Al <sub>2</sub> O <sub>3</sub> , uncoated	44.0	-0.7	unreliable	44	NA	NA
Fe <sub>2</sub> O <sub>3</sub> , uncoated	32.0	-11.0	reliable	-18.1	-9.7	reliable
ZnO, uncoated	71.0	6.5	reliable	-10.8	7.4	reliable
Y <sub>2</sub> O <sub>3</sub> , uncoated	38.0	-11.1	reliable	-10.7	-2.2	reliable
TiO <sub>2</sub> , uncoated	42.3	-11.0	reliable	-9.6	-9.7	reliable
SiO <sub>2</sub> , uncoated	15.0	-0.9	unreliable	-8.1	NA	NA
Bi <sub>2</sub> O <sub>3</sub> , uncoated	90.0	-9.4	reliable	-2.3	7.1	reliable

### 7.3.3 Discussion

The concept of SAPNets and its application on existing nanoinformatics models may support the nanocommunity to go above the current limits of the models and contribute to the nano-data gap-filling. Nevertheless, model combination cannot be performed in an uncontrolled manner, as some possible problems should be considered prior to the mere combination of existing models. To begin with, the models' requirements should be studied, to be used in the appropriate way and derive meaningful predictions (e.g., in this case we tried to reproduce the reported results). Next, the models' “compatibility” should be taken into account which is closely related to the data used to train the separate models. For example, the two models combined here are trained on different data and the descriptor that links them is the zeta potential in water. We tried to ensure that the zeta potential in water is measured under the same conditions and in fact we confirmed that the pH of the water solution in which the zeta potential was measured was in both cases equal to 7. However, we could not corroborate if the rest of the conditions affecting the zeta potential (e.g., the concentration of the NMs in the solution or the temperature) were the same. In addition to the other ambiguities, when testing the SAPNet using the external dataset (

Table 19), we were also unsure for the pH of the solution where the  $\zeta_{H_2O}$  was measured. The data and meta-data sharing for both experimental and computational scientific teams is the only way to ensure transparency at all levels, including the combination of models.

The reliability assessment of the *in silico* models, is the direct comparison of the endpoint predictions with experimental values. In this course, the reliability of the predicted  $\zeta_{KCl}$  values of Table 19 cannot be assessed as there are not available any experimental  $\zeta_{KCl}$  values for this external dataset. In Figure 69, a comparison is made between the predicted  $\zeta_{KCl}$  values generated directly by the MLR model (model B) using experimental  $\zeta_{H_2O}$  values (Table 19) and the  $\zeta_{KCl}$  values generated by the model B of the SAPNet using the predicted  $\zeta_{H_2O}$  input from model A (Table 20). Large differences between the two predicted values are observed, that are possibly due to the differences between  $\zeta_{H_2O}$  experimental and predicted values. It should be noted that to generate the  $\zeta_{H_2O}$  using model A the average diameter of the particles was used in the cases of samples 19 and 22 in Table 19 (the NPs zeta-potential is correlated to their size). In addition, we could not ensure that when training the original-independent models A and B the  $\zeta_{H_2O}$  (which in this SAPNet is used as a linking property) refers to the exact same conditions. Such problems could be addressed by the expansion of the data space, by generating and sharing of data that cover more conditions and can lead to more accurate and reliable predictions. Nonetheless, the concept of SAPNets could be particularly useful when using models that use computational descriptors (they are not based on experimental data). This is the case of the study of Stoliński *et al.*<sup>101</sup> (also presented in § 7.2.2.2 and in Appendix II) where the linking property between the SAPNet models (the nanocluster additive electronegativity) is predicted based on a theoretical model.

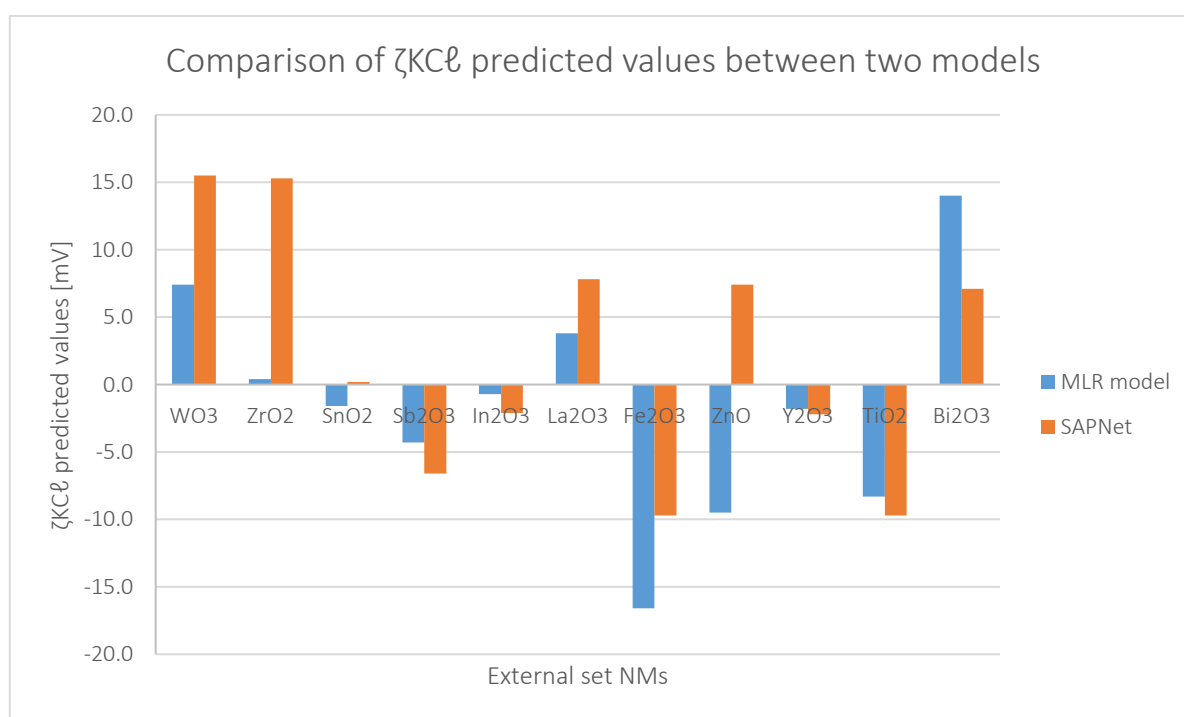


Figure 69: Comparison of the predicted  $\zeta_{KCl}$  values using the Wyrzykowska *et al.*'s MLR model<sup>98</sup> (Table 19) and the SAPNet (Table 20).

The domains of applicability of the models should be also considered and unreliable predictions from the first layers of the SAPNet should be excluded in the next ones. In this case study, both applicability domains are defined and thus, the stakeholders can be aware of the limits of the models and the level of trust to their predictions. This procedure of studying all the aspects of the combined models is a time consuming and tiresome process, and there is a possibility that the stakeholders won't be willing to

spend the required time and/or effort and thus, decide not to use the models to create a SAPNet. For this reason, computational teams should provide all the modelling information in a transparent and understandable manner. In this way, stakeholders will be able to understand if two or more models are suitable to be combined before committing too much time and effort in this process.

Another issue, also raised by Rybińska-Fryca *et al.*<sup>112</sup>, is the possible uncertainty propagation along the network. Certainly, propagation of error occurs since two independent computational studies are combined to retrieve the endpoint values of  $\zeta KCl$ . Here, we made a first effort to enlighten these uncertainties. For a multi-linear equation with the form  $f=f(u,v,w)$ , the error propagation formula contains the partial derivatives of  $f$  with respect to the values of  $u$ ,  $v$  and  $w$  (which are uncorrelated):

$$\delta f = \sqrt{\left(\frac{\partial f}{\partial u} \delta u\right)^2 + \left(\frac{\partial f}{\partial v} \delta v\right)^2 + \left(\frac{\partial f}{\partial w} \delta w\right)^2} \quad [22]$$

Therefore, for the linear regression model  $\zeta KCl = 3.98 + 21.68 \zeta H_2O + 7.88 PN$ , the propagation of error is calculated as follows:

$$\delta(\zeta KCl) = \sqrt{(21.68 \delta(\zeta H_2O))^2 + (7.88 \delta(PN))^2}$$

The error propagation formula considers the uncertainties or errors in the input parameters that are subject to uncertainty or measurement error, which in this case would only be the  $\zeta H_2O$ .  $PN$ , the periodic number, is easily calculated, thus it is not subject to measurement uncertainty.

Then the propagation error is,

$$\delta(\zeta KCl) = 21.68 \delta(\zeta H_2O)$$

According to the principles of error propagation, errors in  $\zeta H_2O$  depend on the uncertainties in the experimental and calculated values. The experimental and calculated values cannot be considered as independent random errors, since only the calculated  $\zeta H_2O$  is used in the SAPNet. In this case, the uncertainty of the experimental  $\zeta H_2O$  value is not relevant in error propagation. The estimation of the uncertainty in the calculated  $\zeta H_2O$  values is based on the principles of error propagation mentioned above and the formula is:

$$\delta(\zeta H_2O_{calc}) = \sqrt{\delta(\zeta H_2O_{exp})^2 + (std\ dev.)^2}$$

For the error in the experimental values of  $\zeta H_2O$ , the assumption is made about the typical magnitude of the uncertainties based on the precision of the measurement. The last decimal place is considered uncertain, therefore  $\delta(\zeta H_2O)_{exp} = 0.1$ . The standard deviation between the experimental and calculated values of zeta potential in water is calculated as follows:

$$std\ dev = \sqrt{\frac{\sum [(\zeta H_2O_{calc,i} - \zeta H_2O_{exp,i}) - avg(\zeta H_2O_{calc} - \zeta H_2O_{exp})]^2}{n}} = 5.46$$

The standard deviation in  $\zeta H_2O$  can be estimated as:

$$\delta(\zeta H_2O_{calc}) = \sqrt{0.1^2 + 5.73^2} = 5.47$$

Eventually, the error propagation in the endpoint  $\zeta KCl$  value is  $\delta(\zeta KCl) = 118.5mV$ . Even though the error that spreads in the final zeta potential in  $KCl$  value is large, it should be noted that in the initial linear regression model, the coefficient of  $\zeta H_2O$  is also high. This means that a relatively small increase in the normalised value of the zeta potential in water results in a great increase in zeta potential in  $KCl$ .

At this point we must mention that the predictions on the  $\zeta_{H_2O}$  derived from model A are at pH 7. Nevertheless, as we may not be sure about the conditions where the respective experimental values of  $\zeta_{H_2O}$  in

Table 19 are acquired, we cannot guarantee that a straightforward comparison of the two values can be made. This is a possible explanation of the large  $\delta(\zeta H_2 O_{calc})$  error that contributes next to the error spread.

## 8 Summary

In this section, the analyses, and the key findings of all the subtasks are gathered, while overall conclusions to drive future developments in the areas of *in silico* approaches for hazard and risk assessment of NMs are presented. Firstly, a search strategy was developed for the literature review of existing computational methods in the nanoinformatics field. This strategy determined the terms of inclusion/exclusion of the available scientific publications and models in the pool of assessed studies. A preliminary assessment of these studies highlighted some shortcomings regarding the availability of datasets for the development of *in silico* methodologies, as well as the use of validation statistics. The key parameters for evaluating the studies were identified, such as the type of employed methodologies, the type of data used for modelling, the documentation of the models, etc.

After that, a further and in-depth analysis of the collected models and tools found in literature and in other reliable sources was performed. Each study was evaluated separately, and information was stored for comparing and assessing methodologies easily. The assessed computational methodologies included nanoQSAR models, grouping/read-across strategies, PBPK models, molecular dynamics simulations, and AOPs. Emphasis was placed on the nanoQSAR and read-across models, including the data used, the applied modelling methodologies, the validation techniques, and the assessment of their availability. Each study was also subjected to a quality review, and the dissemination potential of the various produced models as user-friendly tools was evaluated. The already developed tools were tested and the key parameters of their use were presented in a concise way.

Furthermore, experts in the field of nanoinformatics from Academia and Industry shared their professional opinion regarding the available computational models and methods for NMs hazard and risk assessment, and the potential steps for future improvements. By studying their responses in the respective online survey, we observe that their views converge on some basic issues. One of the main problems that was highlighted is the NMs data scarcity and heterogeneity. Experimental NM data are limited, scattered in different sources (often not accessible for modellers) and formats, and not collected or harmonised under a common protocol. Meta-data regarding the experimental conditions are often also unavailable. This leads to the development of models for a limited number of NMs with limited applicability domains. Another issue that was emphasised is the lack of guidelines regarding the standard properties that should be included in the developed methods by the modellers. From their side, modellers are not always providing proper documentation of their models including all the assumptions or approximations, modelling, validation, and reliability tests that ensure transparency and could reinforce the users' confidence in the models' predictions. To further reinforce the trust in the *in silico* predictions, these should be linked with real cases and applications. Novel techniques should also be developed to address the complexity of NMs, including deep learning models. A communication channel between the stakeholders and the modellers should also be established to exchange feedback and make the *in silico* methods more useful and reliable. Many of the highlighted issues are summarised in one expert's opinion: "There is no open science core (in terms of community standards): no open data core, no open-source core, no open standards core". The lack of community standards and foundations (such as definitions of what can be considered as "Open data", and practices about Open data and data sharing) impedes the integration of the different *in silico* methodologies in routinely-performed risk assessment procedures and thus, the potential of these methods is not fully exploited. FAIR and Open data and models, and in general Open Science could provide a foundation of the knowledge economy.

In compliance with the analysis of the previous findings, we then prepared three case studies for *in silico* models for different applications in NMs of various categories. In the first case study three read-across methods found during the review were employed to assess the anti-microbial interactions

between CNPs (fullerenes, CNTs and GS) and a SARS-CoV-2 RNA fragment, based on physicochemical descriptors. The endpoint of interest was the total potential energy interaction encoding the magnitude of the driving force towards complexation between the CNPs and the cov-RNA fragment. The results of each method were presented in detail and comparisons/consensus results were presented whenever possible. Discussion on the different functionalities of the methods was also performed to highlight their strengths and weaknesses.

In the second case study, seven freely available and user-friendly tools were assessed for predicting toxicities and biodistributions of  $\text{TiO}_2$  and  $\text{TiO}_2$ -based NMs. They all allow stakeholders to perform fast sensitivity analyses or virtually screen newly synthesised NMs, to reduce the time and resources required for experimental evaluation. These tools are suitable for use in an SbD framework to explore the correlations between different properties and the nanotoxicity to define the limits within which NMs are safe and help formulate regulatory limitations. Nonetheless, to guarantee the future viability and use of the tools, frequent feedback is required to create more accurate models as more data appear, and to provide tools with more functions (especially the prediction reliability based on the applicability domain of the background models). Finally, scientific, and computational groups developing such models and tools should invest time in developing comprehensive tutorials that allow stakeholders to exploit the tools to the greatest extent. This should include information on the background methods and training data, input parameters and descriptors, typical results, interpretation of the results, and benchmarking against new experiments. Unless this occurs, the tools will probably not be used widely.

In the third case study, the possibility of the creation of predictive SAPNets was explored, based on models already available in the literature to exploit all the already available models and predict more nanoinformatics-related endpoints. The proposed SAPNet methodology was demonstrated using two consecutive models. The zeta potential in water is predicted from the first model and its value is used as input to the second model to predict the zeta potential in  $\text{KCl}$  solutions. Apart from the demonstration of the methodology, we extensively discussed the problems that we faced when combining different models which have been trained using different datasets. Resolving those issues could lead to trustworthy final predictions that contribute to the SbD of novel NMs or to the fill of data gaps. When researchers aim to combine existing models in the literature to enrich their data, they should be extremely cautious regarding the models' quality and compatibility. Otherwise, predictions derived from problematic or conflicting models, could hold wide margins of error that in turn, makes them practically useless. The descriptors that are the links between the different models should be guaranteed to describe the exact same property (measured under the exact same conditions). The applicability domains of the combined models should be also considered, to ensure the reliability of the predictions across the SAPNet. These uncertainties highlight the need for the development of a communication channel between scientists (e.g., experimentalists, toxicologists, material scientists and computational scientists, etc.) regarding the true nature of the data and what their values represent/express. Furthermore, apart from the data, the meta-data availability and accessibility are equally important, as it is the only way to ensure that meaningful comparisons and combinations can be made and that modelling workflows can be based on sound foundations.

## 8.1 Conclusions and outlook

Most of the collected and evaluated studies are based on solid modelling workflows, algorithms and methodologies and are adequately validated. In addition, the NMs unique characteristics (e.g., the properties of materials at the nanoscale, the NMs dynamic behaviour depending on the environment, etc.) are incorporated in many of the developed methods. Automation and optimisation of the modelling process as well as the incorporation of deep learning methodologies are also considered in

the development of the *in silico* workflows. Progress has been made to create general-use methodologies and algorithms that have been proven to produce reliable predictions even with small datasets. These approaches enable the straightforward development of different models based on different input datasets, allowing their application to various scenarios (endpoints, properties, or types of NMs) without the requirement of a grouping hypothesis. The development of other approaches, such as simulations, PBPK models, and AOPs, that reveal the NMs toxicity mechanisms and the NMs exposure pathways to various organisms can contribute to the SbD of novel NMs. Nonetheless, additional efforts are needed in order to address the limitations that were identified in this project.

The long-standing issue for the development of computational methodologies that will be adequately reliable to support or/and replace standard procedures (often involving experiments on animals) of the NMs risk and hazard assessment is the lack of reliable data and meta-data. Data is needed for the development of nanoQSAR and read-across models and for their statistical validation, but it is also needed to assess their predictive performance in external/real-case scenarios. Data is also necessary to validate other types of methodologies such as PBPKs and MD simulations and to clearly define their applicability domains. In any case, the models' applicability domain clearly defines the data-space areas where predictions can be considered reliable therefore, it strengthens the acceptance and the practical use of the model. A representative sample of the available models was studied during this project. These models were developed and validated -more or less- according to the OECD standards but their actual performance could not be quantified as there is limited data to compare and validate them with. The validation on real-case data was also highlighted by the experts as a barrier towards the solid foundation of confidence and acceptance of the *in silico* testing alternatives from the stakeholders. A possible way to overcome such barriers and to further validate and/or strengthen the stakeholders' confidence, is the post-hoc experimental analysis of the endpoints of interest by comparing the predictions generated by models to the experimental results. Finally, to strengthen the confidence of stakeholders to the *in silico* methodologies transactional research projects can be designed, as proposed by one of the experts. Such projects can bridge the gap between the industry, the academia, and the regulatory bodies, through the co-development of computational methods for NMs risk assessment that will reduce the stakeholders' uncertainties to the outputs of the predictive models. In this course, stakeholders' training can be performed, and user case studies can be also developed to support the integration of the alternative *in silico* methods to the broader nanosafety community.

Data and meta-data FAIRification or Open Data initiatives are important to address the data scarcity and accessibility issues. It is also a step towards data reusability and integration in automated predictive workflows. In this course, universally established test guidelines, ontologies and data harmonisation protocols may contribute to the limitation of data heterogeneity, to provide the community with wide databases of standardised endpoint values for a range of different nanoforms and reliable and accurate experimental data (and meta-data) from well-defined systems. Moreover, the comprehensive data collection is a step towards meta-data availability. For this reason, experimental and computational scientists should be encouraged to use electronic lab notebook (ELN) software to help them document and store the employed methods and protocols, their results in different formats (tables, images, etc.), as well as all the necessary meta-data that are linked to each experiment. The digital format of all this information makes this data searchable from both humans and computers and greatly contributes to data FAIRification and use for modelling purposes. The use of already developed databases such as the eNanoMapper, the NanoPharos, the nanoHUB, etc. can later contribute to the data collection and organisation in a consistent manner.

The inherent complexity of NMs should be also considered for both the development of suitable nanodescriptors (obtained experimentally or theoretically) and of *in silico* methodologies. These novel



descriptors can integrate the multicomponent identification and complexity of the various nanoforms, their different structural and chemical properties, the interactions between NMs and biological environments (e.g., protein corona formations, agglomeration phenomena, etc), as well as the experimental conditions under which different NM properties are measured. Same applies for the methodologies that can integrate the different NM properties and modes-of-action. The NInChI initiative that supports the development of machine-readable text representations of the NMs structure and other properties, may greatly contribute to this nanoinformatics aspect.

Stakeholders may also be reluctant to the computational methodologies, as they are not always properly documented or reported in a comprehensive manner (e.g., including the assumptions made, the modelling input parameters and the expected results explanation, the validation process, etc.). In fact, the analysis of the collected studies revealed certain issues with the use of validation statistics, the accessibility and completeness of datasets for the development of *in silico* approaches, and the proper reporting of the models for reproducibility purposes, using a standardised reported format (e.g., QMRF, MODA). The use of such a reporting template could be a means to overcome these issues enhancing transparency. Following specific guidelines, scientists would be helped to include all the necessary modelling parameters, including dataset details (e.g., the NMs space groups, extra descriptors if calculated, etc.), data manipulation prior to modelling, validation results (including an adequate number of statistical metrics), applicability domain limits, interpretation of the results and availability of the source codes where all the modelling workflow is developed (e.g., through a public repository). Especially when these models are supported by a scientific publication, reviewers should also control and help the authors to improve their work, ensuring that the above objectives are addressed.

During this project, we also encountered issues and errors in some publications that presented models or methodologies. We acknowledge that peer review is a complex process and that reviewers have different perspectives and expertise. However, we believe that there is always room for improvement in the reviewing process to ensure that scientific publications are of the highest quality. Therefore, we suggest that journals take steps to strengthen their reviewing systems to enhance the quality of scientific research. This would help to reduce the amount of noise in literature mining and ensure that scientific research is accurate, reliable, and trustworthy.

Models' dissemination as user-friendly tools is also a step towards their universal integration in the NMs risk assessment activities. More specifically, work has been done to make the models attractive for non-informatics users, and the availability of scripts in open repositories in combination with the availability of deployment platforms will eventually lead to the development of more user-friendly nanoinformatics tools in the future. In this course, we must underline that there is a communication gap between the modelers/tools' developers and the stakeholders. It is often taken for granted that the functionalities of the tool, the type of necessary input data or the expected results are known or obvious. However, this is not always the case as stakeholders come from interdisciplinary fields and may not understand how a tool is used or they may interpret the required input data according to their background, unless a comprehensive manual is available as supporting material of the tool. In addition, the tools should be properly maintained and frequently updated as new data are produced, to ensure their permanence in the future. This means that stakeholders who invest (effort, time, etc.) in an *in silico* tool should be able to maintain their trust for a long-term period. Finally, server connectivity issues should be also addressed in order to support the usability of the tools.

The Transparency, Reliability, Accessibility, Applicability and Completeness (TRAAC framework)<sup>3</sup> can be applied to evaluate the *in silico* tools quality in terms of compatibility with regulatory frameworks and usefulness for end-users. This framework through the assessment and scoring of the tools can quantify

their readiness towards their wider regulatory acceptance, as well as it can identify which barriers impede the user-acceptance, based on a series of pre-defined criteria.

Dynamic communication and collaboration between the different researchers, stakeholders and regulatory experts is crucial for the development of robust, reliable, and meaningful descriptors and computational models. Data and knowledge can be transferred between interdisciplinary groups and thus, each expert can contribute to a different aspect of the nanoinformatics applications enhancing in that way the usefulness and applicability of computational methodologies. Such synergies will eventually contribute to the accomplishment of the “Closer-to-the-Market-Roadmap” recommendations as established by the European NanoSafety Cluster, for the market implementation of safe NMs and nano-enabled products<sup>221</sup>.

## Abbreviations

- ADME: absorption, distribution, metabolism, and excretion
- AI: Artificial intelligence
- AIC: Akaike information criterion
- ANN: Artificial neural networks
- AO: Adverse outcome
- AOP: Adverse outcome pathway
- API: Application programming interface
- AUC: Area under the curve
- ATP: Adenosine triphosphate
- BMD: Benchmark dose
- BMR: Basal metabolic rate
- BN: Bayesian network
- BSA: Bovine serum albumin
- CCC: Concordance correlation coefficient
- CG: Coarse-grained
- CNP: Carbon-based nanoparticles
- CNT: Carbon nanotube
- CSV: Comma-separated values
- DoA: Domain of applicability
- $E_{\text{int}}$ : Total potential energy interaction
- $EC_{50}$ : Median effective concentration
- ECHA: European Chemicals Agency
- ELN: Electronic lab notebook
- EMMC: European Materials Modelling Council
- EPA: U.S. Environmental Protection Agency
- EU: European Union
- EUON: European Union Observatory of nanomaterials
- FAIR (data, models): Findable, accessible, interoperable, and reusable
- FF: Far field
- FN: False negative
- FP: False positive
- FPR: False positive rate
- GA: Genetic algorithm
- GO: Graphene oxide
- GS: Graphene sheet
- GUI: Graphical user interface

- HARM: High aspect ratio material
- HTS: High throughput screening
- IATA: Integrated approach to testing and assessment
- ICH: International council for harmonisation of technical requirements for pharmaceuticals for human use
- ICRP: International commission on radiological protection
- InChI: International Chemical Identifier
- JRC: Joint Research Centre
- KE: Key event
- KER: Key event relationship
- $k$ NN:  $k$ -nearest neighbour
- LC<sub>50</sub>: Median lethal concentration
- LDA: Linear discriminant analysis
- LDH: Lactate Dehydrogenase
- LF: Lung fibrosis
- LOO: Leave-one-out
- MAE: Mean absolute error
- MCC: Matthews' correlation coefficient
- MD: Molecular dynamics
- MeOx: Metal oxide
- MIE: Molecular initiating event
- MILP: Mixed integer linear programming
- ML: Machine learning
- MLR: Multiple linear regression
- MoA: Mechanism of action
- MODA: Modelling data (reporting template)
- MPPD: Multiple-path particle dosimetry (model)
- MSE: Mean squared error
- MW: Molecular weight
- MWCNT: Multi-walled carbon nanotube
- NF: Near field
- NIA: Nanotechnology Industries Association
- NM: Nanomaterial
- NP: Nanoparticle
- OECD: Organisation for economic co-operation and development
- OEL: Occupational exposure limit
- OPLS: Orthogonal partial least squares
- OSA: Overall surface area
- PAA: Polyacrilamide
- PBPK: Physiologically-based pharmacokinetics (model)
- PCA: Principal component analysis
- pEC<sub>50</sub>: Negative logarithm of the EC<sub>50</sub>
- PLS: Partial least squares
- PN: Periodic number
- POD: Point of departure
- QC: Quality control
- QMRF: QSAR model reporting format
- QNAR: Quantitative nanostructure-activity relationship
- QPRF: QSAR prediction reporting format

- QSAR: Quantitative structure-activity relationship
- QSPR: Quantitative structure-property relationship
- QSTR: Quantitative structure-toxicity relationship
- RAAF: Read-across assessment framework
- RASAR: Read-across structure–activity relationships
- REACH: Registration, evaluation, authorization, and restriction of chemicals
- RMSE: Root mean square error
- RMESP: Root mean square error of prediction
- ROC: Receiver operating characteristic
- SAPNets: Structure-activity prediction networks
- SbD: Safety-by-design
- SDeg: Sum of degrees of every carbon atom
- SOM: Self organising map
- SPION: Superparamagnetic iron oxide nanoparticle
- SSA: Specific surface area
- SVM: Support vector machines
- TN: True negative
- TNR: True negative rate
- TP: True positive
- TPR: True positive rate
- TRAAC: Transparency, reliability, accessibility, applicability and completeness
- UV: Ultraviolet
- WoE: Weight of evidence
- XGB: Extreme gradient boosting

## References

1. Papadiamantis, A. G. *et al.* Computational enrichment of physicochemical data for the development of a  $\zeta$ -potential read-across predictive model with Isalos Analytics Platform. *NanoImpact* **22**, 100308 (2021).
2. Lynch, I. *et al.* Can an InChI for nano address the need for a simplified representation of complex nanomaterials across experimental and nanoinformatics studies? *Nanomaterials* **10**, 2493 (2020).
3. Shandilya, N. *et al.* TRAAC framework for regulatory acceptance and wider usability of tools and methods for safe innovation and sustainability of manufactured nanomaterials. *Res. Sq.* **30**, 100461 (2021).
4. Nature. Reporting standards and availability of data, materials, code and protocols. <https://www.nature.com/nature-portfolio/editorial-policies/reporting-standards#availability-of-computer-code>.
5. Hubrecht, R. C. & Carter, E. The 3Rs and humane experimental technique: Implementing change. *Animals* **9**, 1–10 (2019).
6. Forest, V. Experimental and Computational Nanotoxicology—Complementary Approaches for Nanomaterial Hazard Assessment. *Nanomaterials* **12**, 1346 (2022).
7. Buglak, A. A., Zherdev, A. V. & Dzantiev, B. B. Nano-(Q)SAR for cytotoxicity prediction of engineered nanomaterials. *Molecules* vol. 24 at <https://doi.org/10.3390/molecules24244537> (2019).
8. von Ranke, N. L. *et al.* Applying in silico approaches to nanotoxicology: Current status and future

- potential. *Comput. Toxicol.* **22**, 100225 (2022).
9. Basei, G. *et al.* Making use of available and emerging data to predict the hazards of engineered nanomaterials by means of in silico tools: A critical review. *NanoImpact* **13**, 76–99 (2019).
  10. Oomen, A. *et al.* Grouping and Read-Across Approaches for Risk Assessment of Nanomaterials. *Int. J. Environ. Res. Public Health* **12**, 13415–13434 (2015).
  11. European Chemicals Agency & ECHA. *Read-Across Assessment Framework (RAAF)*. European Chemicals Agency (2017). doi:10.2823/619212.
  12. European Chemicals Agency. *Appendix R . 6-1: Recommendations for nanomaterials applicable to the Guidance on QSARs and Grouping*. [https://echa.europa.eu/documents/10162/23036412/appendix\\_r6\\_nanomaterials\\_en.pdf/71ad76f0-ab4c-fb04-acba-074cf045eaaa](https://echa.europa.eu/documents/10162/23036412/appendix_r6_nanomaterials_en.pdf/71ad76f0-ab4c-fb04-acba-074cf045eaaa) (2019) doi:10.2823/273911.
  13. Liang, X. *et al.* Physiologically Based Pharmacokinetic Model for Long-Circulating Inorganic Nanoparticles. *Nano Lett.* **16**, 939–945 (2016).
  14. Bachler, G., Von Goetz, N. & Hungerbuhler, K. Using physiologically based pharmacokinetic (PBPK) modeling for dietary risk assessment of titanium dioxide (TiO<sub>2</sub>) nanoparticles. *Nanotoxicology* **9**, 373–380 (2015).
  15. Tsiros, P. *et al.* Towards an in silico integrated approach for testing and assessment of nanomaterials: from predicted indoor air concentrations to lung dose and biodistribution. *Environ. Sci. Nano* **9**, 1282–1297 (2022).
  16. Li, M., Al-Jamal, K. T., Kostarelos, K. & Reineke, J. Physiologically based pharmacokinetic modeling of nanoparticles. *ACS Nano* **4**, 6303–6317 (2010).
  17. Utembe, W., Clewell, H. J., Sanabria, N., Doganis, P. & Gulumian, M. Current approaches and techniques in physiologically based pharmacokinetic (PBPK) modelling of nanomaterials. *Nanomaterials* **10**, 1–32 (2020).
  18. Pem, B. *et al.* Fate and Transformation of Silver Nanoparticles in Different Biological Conditions. *Beilstein J. Nanotechnol.* **12**, 665–679 (2021).
  19. Klaessig, F. C. PBPK Modeling of Slightly Soluble Silver Nanomaterials and Regulatory Acceptance. *Small* **16**, 1–8 (2020).
  20. Winkler, D. A. Recent advances, and unresolved issues, in the application of computational modelling to the prediction of the biological effects of nanomaterials. *Toxicol. Appl. Pharmacol.* **299**, 96–100 (2016).
  21. Afantitis, A. *et al.* NanoSolveIT Project: Driving nanoinformatics research to develop innovative and integrated tools for in silico nanosafety assessment. *Comput. Struct. Biotechnol. J.* **18**, 583–602 (2020).
  22. Guo, Z. *et al.* Surface Functionalization of Graphene-Based Materials: Biological Behavior, Toxicology, and Safe-By-Design Aspects. *Adv. Biol.* **5**, (2021).
  23. Gerloff, K. *et al.* The Adverse Outcome Pathway approach in nanotoxicology. *Comput. Toxicol.* **1**, 3–11 (2017).
  24. Society for the Advancement of Adverse Outcome Pathways. AOP-Wiki. <https://aopwiki.org/>.
  25. Organisation for Economic Co-operation and Development. Users Handbook supplement to the Guidance Document for developing and assessing Adverse Outcome Pathways. *OECD Ser.*

*Advers. Outcome Pathways* (2018) doi:10.1787/5jlv1m9d1g32-en.

26. Gajewicz, A., Cronin, M. T. D., Rasulev, B., Leszczynski, J. & Puzyn, T. Novel approach for efficient predictions properties of large pool of nanomaterials based on limited set of species: Nano-read-across. *Nanotechnology* **26**, 15701 (2015).
27. Giusti, A. *et al.* Nanomaterial grouping: Existing approaches and future recommendations. *NanoImpact* **16**, 100182 (2019).
28. Li, M., Zou, P., Tyner, K. & Lee, S. Physiologically Based Pharmacokinetic (PBPK) Modeling of Pharmaceutical Nanoparticles. *AAPS J.* **19**, 26–42 (2017).
29. Minnema, J., Borgos, S. E. F., Liptrott, N., Vandebriel, R. & Delmaar, C. Physiologically based pharmacokinetic modeling of intravenously administered nanoformulated substances. *Drug Deliv. Transl. Res.* **12**, 2132–2144 (2022).
30. Leist, M. *et al.* Adverse outcome pathways: opportunities, limitations and open questions. *Arch. Toxicol.* **91**, 3477–3505 (2017).
31. Halappanavar, S., Ede, J. D., Shatkin, J. A. & Krug, H. F. A systematic process for identifying key events for advancing the development of nanomaterial relevant adverse outcome pathways. *NanoImpact* **15**, 100178 (2019).
32. Furxhi, I. Health and environmental safety of nanomaterials: O Data, Where Art Thou? *NanoImpact* **25**, 100378 (2022).
33. Walkey, C. D. *et al.* Protein Corona Fingerprinting Predicts the Cellular Interaction of Gold and Silver Nanoparticles. *ACS Nano* **8**, 2439–2455 (2014).
34. Varsou, D.-D. & Sarimveis, H. Apellis: An online tool for read-across model development. *Comput. Toxicol.* **17**, 100146 (2021).
35. Varsou, D.-D., Koutroumpa, N.-M. & Sarimveis, H. Automated Grouping of Nanomaterials and Read-Across Prediction of Their Adverse Effects Based on Mathematical Optimization. *J. Chem. Inf. Model.* **61**, 2766–2779 (2021).
36. Afantitis, A., Melagraki, G., Tsoumanis, A., Valsami-Jones, E. & Lynch, I. A nanoinformatics decision support tool for the virtual screening of gold nanoparticle cellular association using protein corona fingerprints. *Nanotoxicology* (2018) doi:10.1080/17435390.2018.1504998.
37. Varsou, D.-D. *et al.* toxFlow: A Web-Based Application for Read-Across Toxicity Prediction Using Omics and Physicochemical Data. *J. Chem. Inf. Model.* **58**, 543–549 (2018).
38. Helma, C., Rautenberg, M. & Gebele, D. Nano-Lazar: Read across Predictions for Nanoparticle Toxicities with Calculated and Measured Properties. *Front. Pharmacol.* **8**, 1–11 (2017).
39. Tsiliki, G. *et al.* RRegrs: An R package for computer-aided model selection with multiple regression models. *J. Cheminform.* **7**, (2015).
40. Liu, R., Jiang, W., Walkey, C. D., Chan, W. C. W. & Cohen, Y. Prediction of nanoparticles-cell association based on corona proteins and physicochemical properties. *Nanoscale* **7**, 9664–9675 (2015).
41. Papa, E., Doucet, J. P., Sangion, A. & Doucet-Panaye, A. Investigation of the influence of protein corona composition on gold nanoparticle bioactivity using machine learning approaches. *SAR QSAR Environ. Res.* **27**, 521–538 (2016).
42. Weissleder, R., Kelly, K., Sun, E. Y., Shtatland, T. & Josephson, L. Cell-specific targeting of

- nanoparticles by multivalent attachment of small molecules. *Nat. Biotechnol.* **23**, 1418–1423 (2005).
43. Fourches, D. *et al.* Quantitative nanostructure - Activity relationship modeling. *ACS Nano* **4**, 5703–5712 (2010).
  44. Melagraki, G. & Afantitis, A. Enalos InSilicoNano platform: an online decision support tool for the design and virtual screening of nanoparticles. *RSC Adv.* **4**, 50713–50725 (2014).
  45. Kar, S., Gajewicz, A., Puzyn, T. & Roy, K. Nano-quantitative structure-activity relationship modeling using easily computable and interpretable descriptors for uptake of magnetofluorescent engineered nanoparticles in pancreatic cancer cells. *Toxicol. Vitr.* **28**, 600–606 (2014).
  46. Ghorbanzadeh, M., Fatemi, M. H. & Karimpour, M. Modeling the Cellular Uptake of Magnetofluorescent Nanoparticles in Pancreatic Cancer Cells: A Quantitative Structure Activity Relationship Study. *Ind. Eng. Chem. Res.* **51**, 10712–10718 (2012).
  47. Chau, Y. T. & Yap, C. W. Quantitative Nanostructure–Activity Relationship modelling of nanoparticles. *RSC Adv.* **2**, 8489 (2012).
  48. Epa, V. C. *et al.* Modelling biological activities of nanoparticles. *Nano Lett.* **12**, 1–7 (2012).
  49. Oksel, C., Winkler, D. A., Ma, C. Y., Wilkins, T. & Wang, X. Z. Accurate and interpretable nanoSAR models from genetic programming-based decision tree construction approaches. *Nanotoxicology* **10**, 1001–1012 (2016).
  50. Puzyn, T. *et al.* Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles. *Nat. Nanotechnol.* **6**, 175–178 (2011).
  51. Kar, S., Gajewicz, A., Roy, K., Leszczynski, J. & Puzyn, T. Extrapolating between toxicity endpoints of metal oxide nanoparticles: Predicting toxicity to Escherichia coli and human keratinocyte cell line (HaCaT) with Nano-QTTR. *Ecotoxicol. Environ. Saf.* **126**, 238–244 (2016).
  52. Kar, S., Pathakoti, K., Tchounwou, P. B., Leszczynska, D. & Leszczynski, J. Evaluating the cytotoxicity of a large pool of metal oxide nanoparticles to Escherichia coli: Mechanistic understanding through In Vitro and In Silico studies. *Chemosphere* **264**, 128428 (2021).
  53. Kar, S., Gajewicz, A., Puzyn, T., Roy, K. & Leszczynski, J. Periodic table-based descriptors to encode cytotoxicity profile of metal oxide nanoparticles: A mechanistic QSTR approach. *Ecotoxicol. Environ. Saf.* **107**, 162–169 (2014).
  54. Pan, Y. *et al.* Nano-QSAR modeling for predicting the cytotoxicity of metal oxide nanoparticles using novel descriptors. *RSC Adv.* **6**, 25766–25775 (2016).
  55. Gajewicz, A. What if the number of nanotoxicity data is too small for developing predictive Nano-QSAR models? An alternative read-across based approach for filling data gaps. *Nanoscale* **9**, 8435–8448 (2017).
  56. Sizochenko, N. *et al.* How the toxicity of nanomaterials towards different species could be simultaneously evaluated: a novel multi-nano-read-across approach. *Nanoscale* **10**, 582–591 (2018).
  57. Gajewicz, A., Jagiello, K., Cronin, M. T. D. D., Leszczynski, J. & Puzyn, T. Addressing a bottle neck for regulation of nanomaterials: quantitative read-across (Nano-QRA) algorithm for cases when only limited data is available. *Environ. Sci. Nano* **4**, 346–358 (2017).
  58. Toropov, A. A. *et al.* Novel application of the CORAL software to model cytotoxicity of metal

- oxide nanoparticles to bacteria *Escherichia coli*. *Chemosphere* **89**, 1098–1102 (2012).
59. Chatterjee, M., Banerjee, A., De, P., Gajewicz-Skretna, A. & Roy, K. A novel quantitative read-across tool designed purposefully to fill the existing gaps in nanosafety data. *Environ. Sci. Nano* **9**, 189–203 (2022).
  60. Mu, Y. *et al.* Predicting toxic potencies of metal oxide nanoparticles by means of nano-QSARs. *Nanotoxicology* **10**, 1207–1214 (2016).
  61. Basant, N. & Gupta, S. Multi-target QSTR modeling for simultaneous prediction of multiple toxicity endpoints of nano-metal oxides. *Nanotoxicology* **11**, 339–350 (2017).
  62. De, P., Kar, S., Roy, K. & Leszczynski, J. Second generation periodic table-based descriptors to encode toxicity of metal oxide nanoparticles to multiple species: QSTR modeling for exploration of toxicity mechanisms. *Environ. Sci. Nano* **5**, 2742–2760 (2018).
  63. Gajewicz, A. *et al.* Towards understanding mechanisms governing cytotoxicity of metal oxides nanoparticles: Hints from nano-QSAR studies. *Nanotoxicology* **9**, 313–325 (2015).
  64. Shaw, S. Y. *et al.* Perturbational profiling of nanomaterial biologic activity. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 7387–7392 (2008).
  65. Melagraki, G. & Afantitis, A. A Risk Assessment Tool for the Virtual Screening of Metal Oxide Nanoparticles through Enalos InSilicoNano Platform. *Curr. Top. Med. Chem.* **15**, 1827–1836 (2015).
  66. Liu, R. *et al.* Nano-SAR development for bioactivity of nanoparticles with considerations of decision boundaries. *Small* **9**, 1842–1852 (2013).
  67. Zhou, H. *et al.* A nano-combinatorial library strategy for the discovery of nanotubes with reduced protein-binding, cytotoxicity, and immune response. *Nano Lett.* **8**, 859–865 (2008).
  68. Fourches, D. *et al.* Computer-aided design of carbon nanotubes with the desired bioactivity and safety profiles. *Nanotoxicology* **10**, 374–383 (2016).
  69. Shao, C.-Y. *et al.* Dependence of QSAR Models on the Selection of Trial Descriptor Sets: A Demonstration Using Nanotoxicity Endpoints of Decorated Nanotubes. *J. Chem. Inf. Model.* **53**, 142–158 (2013).
  70. Varsou, D.-D. *et al.* A safe-by-design tool for functionalised nanomaterials through the Enalos Nanoinformatics Cloud platform. *Nanoscale Adv.* **1**, 706–718 (2019).
  71. Oksel, C., Ma, C. Y. & Wang, X. Z. Current situation on the availability of nanostructure–biological activity data. *SAR QSAR Environ. Res.* **26**, 79–94 (2015).
  72. Robinson, R. L. M. *et al.* Identifying diverse metal oxide nanomaterials with lethal effects on embryonic zebrafish using machine learning. *Beilstein J. Nanotechnol.* **12**, 1297–1325 (2021).
  73. Wang, X., Liu, L., Zhang, W. & Ma, X. Prediction of Plant Uptake and Translocation of Engineered Metallic Nanoparticles by Machine Learning. *Environ. Sci. Technol.* **55**, 7491–7500 (2021).
  74. Jung, U., Lee, B., Kim, G., Shin, H. K. & Kim, K. T. Nano-QTTR development for interspecies aquatic toxicity of silver nanoparticles between daphnia and fish. *Chemosphere* **283**, 131164 (2021).
  75. Varsou, D.-D., Ellis, L.-J. A., Afantitis, A., Melagraki, G. & Lynch, I. Ecotoxicological read-across models for predicting acute toxicity of freshly dispersed versus medium-aged NMs to *Daphnia magna*. *Chemosphere* **285**, 131452 (2021).
  76. Kotzabasaki, M. I., Sotiropoulos, I. & Sarimveis, H. QSAR modeling of the toxicity classification of



- superparamagnetic iron oxide nanoparticles (SPIONs) in stem-cell monitoring applications: An integrated study from data curation to model development. *RSC Adv.* **10**, 5385–5391 (2020).
77. Kotzabasaki, M., Sotiropoulos, I., Charitidis, C. & Sarimveis, H. Machine learning methods for multi-walled carbon nanotubes (MWCNT) genotoxicity prediction. *Nanoscale Adv.* **3**, 3167–3176 (2021).
  78. Zhang, H. *et al.* Use of Metal Oxide Nanoparticle Band Gap to Develop a Predictive Paradigm for Oxidative Stress and Acute Pulmonary Inflammation. *ACS Nano* **6**, 4349–4368 (2012).
  79. Bigdeli, A., Hormozi-Nezhad, M. R. & Parastar, H. Using nano-QSAR to determine the most responsible factor(s) in gold nanoparticle exocytosis. *RSC Adv.* **5**, 57030–57037 (2015).
  80. Sizochenko, N., Mikolajczyk, A., Syzochenko, M., Puzyn, T. & Leszczynski, J. Zeta potentials ( $\zeta$ ) of metal oxide nanoparticles: A meta-analysis of experimental data and a predictive neural networks modeling. *NanoImpact* **22**, 100317 (2021).
  81. Vijayaraghavan, V., Garg, A., Wong, C. H., Tai, K. & Bhalerao, Y. Predicting the mechanical characteristics of hydrogen functionalized graphene sheets using artificial neural network approach. *J. Nanostructure Chem.* **3**, (2013).
  82. Wang, Q., Apul, O. G., Xuan, P., Luo, F. & Karanfil, T. Development of a 3D QSPR model for adsorption of aromatic compounds by carbon nanotubes: Comparison of multiple linear regression, artificial neural network and support vector machine. *RSC Adv.* **3**, 23924–23934 (2013).
  83. Fernandez, M., Shi, H. & Barnard, A. S. Quantitative Structure-Property Relationship Modeling of Electronic Properties of Graphene Using Atomic Radial Distribution Function Scores. *J. Chem. Inf. Model.* **55**, 2500–2506 (2015).
  84. Findlay, M. R., Freitas, D. N., Mobed-Miremadi, M. & Wheeler, K. E. Machine learning provides predictive analysis into silver nanoparticle protein corona formation from physicochemical properties. *Environ. Sci. Nano* **5**, 64–71 (2018).
  85. Gajewicz, A. *et al.* Decision tree models to classify nanomaterials according to the DF4nanoGrouping scheme. *Nanotoxicology* **12**, 1–17 (2018).
  86. Witten Ian H and Frank, Eibe and Hall, Mark A and Pal, C. J. *Data Mining: Practical Machine Learning Tools and Techniques*. (Morgan Kaufmann, 2016). doi:10.1016/C2009-0-19715-5.
  87. Li, J. *et al.* Feature Selection: A Data Perspective. *ACM Comput. Surv.* **50**, 1–45 (2018).
  88. Thwala, M. M. *et al.* Using the Isalos platform to develop a (Q)SAR model that predicts metal oxide toxicity utilizing facet-based electronic, image analysis-based, and periodic table derived properties as descriptors. *Struct. Chem.* **33**, 527–538 (2022).
  89. Varsou, D. *et al.* Zeta-Potential Read-Across Model Utilizing Nanodescriptors Extracted via the NanoXtract Image Analysis Tool Available on the Enalos Nanoinformatics Cloud Platform. *Small* **16**, 1906588 (2020).
  90. Sang, L. *et al.* Machine Learning for Evaluating the Cytotoxicity of Mixtures of Nano-TiO<sub>2</sub> and Heavy Metals: QSAR Model Apply Random Forest Algorithm after Clustering Analysis. *Molecules* **27**, 6125 (2022).
  91. Cai, Y., Nowack, B. & Wigger, H. Identifying ecotoxicological descriptors to enable predictive hazard assessments of nano-TiO<sub>2</sub> from a meta-analysis of ecotoxicological data. *NanoImpact* **15**, (2019).

92. Qi, R. *et al.* Toward comprehension of the cytotoxicity of heterogeneous TiO<sub>2</sub>-based engineered nanoparticles: a nano-QSAR approach. *Environ. Sci. Nano* **8**, 927–936 (2021).
93. George, S. *et al.* Use of a high-throughput screening approach coupled with in vivo zebrafish embryo screening to develop hazard ranking for engineered nanomaterials. *ACS Nano* **5**, 1805–1817 (2011).
94. Burello, E. & Worth, A. P. A theoretical framework for predicting the oxidative stress potential of oxide nanoparticles. *Nanotoxicology* **5**, 228–235 (2011).
95. Furxhi, I., Murphy, F., Mullins, M. & Poland, C. A. Machine learning prediction of nanoparticle in vitro toxicity: A comparative study of classifiers and ensemble-classifiers using the Copeland Index. *Toxicol. Lett.* **312**, 157–166 (2019).
96. Luan, F. *et al.* Computer-aided nanotoxicology: Assessing cytotoxicity of nanoparticles under diverse experimental conditions by using a novel QSTR-perturbation approach. *Nanoscale* **6**, 10623–10630 (2014).
97. Yan, X., Zhang, J., Russo, D. P., Zhu, H. & Yan, B. Prediction of Nano-Bio Interactions through Convolutional Neural Network Analysis of Nanostructure Images. *ACS Sustain. Chem. Eng.* **8**, 19096–19104 (2020).
98. Wyrzykowska, E., Mikolajczyk, A., Sikorska, C. & Puzyn, T. Development of a novel in silico model of zeta potential for metal oxide nanoparticles: A nano-QSPR approach. *Nanotechnology* **27**, 1–8 (2016).
99. Karatzas, P. *et al.* Development of Deep Learning Models for Predicting the Effects of Exposure to Engineered Nanomaterials on *Daphnia magna*. *Small* **16**, (2020).
100. Mikolajczyk, A. *et al.* A chemoinformatics approach for the characterization of hybrid nanomaterials: Safer and efficient design perspective. *Nanoscale* **11**, 11808–11818 (2019).
101. Stoliński, F., Rybińska-Fryca, A., Gromelski, M., Mikolajczyk, A. & Puzyn, T. NanoMixHamster: a web-based tool for predicting cytotoxicity of TiO<sub>2</sub>-based multicomponent nanomaterials toward Chinese hamster ovary (CHO-K1) cells. *Nanotoxicology* **16**, 276–289 (2022).
102. Aschberger, K. *et al.* Grouping of multi-walled carbon nanotubes to read-across genotoxicity: A case study to evaluate the applicability of regulatory guidance. *Comput. Toxicol.* **9**, 22–35 (2019).
103. Swirog, M. *et al.* Predicting electrophoretic mobility of TiO<sub>2</sub>, ZnO, and CeO<sub>2</sub> nanoparticles in natural waters: The importance of environment descriptors in nanoinformatics models. *Sci. Total Environ.* **840**, 0–20 (2022).
104. Forest, V. *et al.* Towards an alternative to nano-QSAR for nanoparticle toxicity ranking in case of small datasets. *J. Nanoparticle Res.* **21**, (2019).
105. Oksel, C., Ma, C. Y. & Wang, X. Z. Structure-activity relationship models for hazard assessment and risk management of engineered nanomaterials. in *Procedia Engineering* vol. 102 1500–1510 (Elsevier Ltd, 2015).
106. Liu, R. *et al.* Association rule mining of cellular responses induced by metal and metal oxide nanoparticles. *Analyst* **139**, 943–953 (2014).
107. Wang, W. *et al.* Predicting Nano–Bio Interactions by Integrating Nanoparticle Libraries and Quantitative Nanostructure Activity Relationship Modeling. *ACS Nano* **11**, 12641–12649 (2017).
108. Kaweeteerawat, C. *et al.* Toxicity of Metal Oxide Nanoparticles in *Escherichia coli* Correlates with Conduction Band and Hydration Energies. *Environ. Sci. Technol.* **49**, 1105–1112 (2015).

109. Romeo, D., Louka, C., Gudino, B., Wigström, J. & Wick, P. Structure-activity relationship of graphene-related materials: A meta-analysis based on mammalian in vitro toxicity data. *NanoImpact* **28**, 100436 (2022).
110. Liu, R. *et al.* Classification NanoSAR Development for Cytotoxicity of Metal Oxide Nanoparticles. *Small* **7**, 1118–1126 (2011).
111. Jha, S. K. & Yoon, T. H. Toxicity modelling of nanomaterials by origin evaluation of their physicochemical descriptors using a combination of principal component analysis and support vector machine methods. *Expert Syst.* **37**, 1–14 (2020).
112. Rybińska-Fryca, A., Mikolajczyk, A. & Puzyn, T. Structure-activity prediction networks (SAPNets): A step beyond Nano-QSAR for effective implementation of the safe-by-design concept. *Nanoscale* **12**, 20669–20676 (2020).
113. Lamon, L. *et al.* Grouping of nanomaterials to read-across hazard endpoints: From data collection to assessment of the grouping hypothesis by application of chemoinformatic techniques. *Part. Fibre Toxicol.* **15**, 1–17 (2018).
114. Bahl, A. *et al.* Recursive feature elimination in random forest classification supports nanomaterial grouping. *NanoImpact* **15**, 100179 (2019).
115. Alsharif, S. A., Power, D., Rouse, I. & Lobaskin, V. In silico prediction of protein adsorption energy on titanium dioxide and gold nanoparticles. *Nanomaterials* **10**, 1–21 (2020).
116. Lazarovits, J. *et al.* Supervised Learning and Mass Spectrometry Predicts the in Vivo Fate of Nanomaterials. *ACS Nano* **13**, 8023–8034 (2019).
117. Gousiadou, C. *et al.* Machine learning predictions of concentration-specific aggregate hazard scores of inorganic nanomaterials in embryonic zebrafish. *Nanotoxicology* **15**, 446–476 (2021).
118. Balraadsing, S., Peijnenburg, W. J. G. M. & Vijver, M. G. Exploring the potential of in silico machine learning tools for the prediction of acute *Daphnia magna* nanotoxicity. *Chemosphere* **307**, 135930 (2022).
119. OECD & Organisation for Economic Co-operation and Development. Guidance document on the validation of (quantitative) structure-activity relationship [(Q)SAR] models. 1–154 (2007).
120. Melagraki, G. & Afantitis, A. Enalos InSilicoNano platform: An online decision support tool for the design and virtual screening of nanoparticles. *RSC Adv.* **4**, 50713–50725 (2014).
121. Kleandrova, V. V. *et al.* Computational Tool for Risk Assessment of Nanomaterials: Novel QSTR-Perturbation Model for Simultaneous Prediction of Ecotoxicity and Cytotoxicity of Uncoated and Coated Nanoparticles under Multiple Experimental Conditions. *Environ. Sci. Technol.* **48**, 14686–14694 (2014).
122. Singh, K. P. & Gupta, S. Nano-QSAR modeling for predicting biological activity of diverse nanomaterials. *RSC Adv.* **4**, 13215–13230 (2014).
123. Varsou, D.-D., Afantitis, A., Melagraki, G. & Sarimveis, H. Read-across predictions of nanoparticle hazard endpoints: a mathematical optimization approach. *Nanoscale Adv.* **1**, 3485–3498 (2019).
124. Shayanfar, S. & Shayanfar, A. Comparison of various methods for validity evaluation of QSAR models. *BMC Chem.* **16**, 1–9 (2022).
125. Arts, J. H. E. *et al.* A decision-making framework for the grouping and testing of nanomaterials (DF4nanoGrouping). *Regul. Toxicol. Pharmacol.* **71**, S1–S27 (2015).

126. El Yamani, N. *et al.* Hazard identification of nanomaterials: In silico unraveling of descriptors for cytotoxicity and genotoxicity. *Nano Today* **46**, 101581 (2022).
127. Zhang, F., Wang, Z., Vijver, M. G. & Peijnenburg, W. J. G. M. Probing nano-QSAR to assess the interactions between carbon nanoparticles and a SARS-CoV-2 RNA fragment. *Ecotoxicol. Environ. Saf.* **219**, 112357 (2021).
128. Ramezanizadeh, M. & Nazari, M. A. Modeling thermal conductivity of Ag/water nanofluid by applying a mathematical correlation and artificial neural network. *Int. J. Low-Carbon Technol.* **14**, 468–474 (2019).
129. Hassanzadeh, Z., Kompany-Zareh, M., Ghavami, R., Gholami, S. & Malek-Khatabi, A. Combining radial basis function neural network with genetic algorithm to QSPR modeling of adsorption on multi-walled carbon nanotubes surface. *J. Mol. Struct.* **1098**, 191–198 (2015).
130. Trinh, T. X., Seo, M., Yoon, T. H. & Kim, J. Developing random forest based QSAR models for predicting the mixture toxicity of TiO<sub>2</sub> based nano-mixtures to *Daphnia magna*. *NanoImpact* **25**, 100383 (2022).
131. Manganelli, S., Leone, C., Toropov, A. A., Toropova, A. P. & Benfenati, E. QSAR model for predicting cell viability of human embryonic kidney cells exposed to SiO<sub>2</sub> nanoparticles. *Chemosphere* **144**, 995–1001 (2016).
132. Toropov, A. A. & Toropova, A. P. Quasi-QSAR for mutagenic potential of multi-walled carbon-nanotubes. *Chemosphere* **124**, 40–46 (2015).
133. Subramanian, N. A. & Palaniappan, A. NanoTox: Development of a Parsimonious in Silico Model for Toxicity Assessment of Metal-Oxide Nanoparticles Using Physicochemical Features. *ACS Omega* **6**, 11729–11739 (2021).
134. Sizochenko, N. *et al.* Causal inference methods to assist in mechanistic interpretation of classification nano-SAR models. *RSC Adv.* **5**, 77739–77745 (2015).
135. Furxhi, I. *et al.* Application of Bayesian networks in determining nanoparticle-induced cellular outcomes using transcriptomics. *Nanotoxicology* **13**, 827–848 (2019).
136. Naser, M. Z. & Alavi, A. H. Error Metrics and Performance Fitness Indicators for Artificial Intelligence and Machine Learning in Engineering and Sciences. *Archit. Struct. Constr.* (2021) doi:10.1007/s44150-021-00015-8.
137. Denis Fourches, Dongqiuye Pu, Carlos Tassa, Ralph Weissleder, Stanley Y. Shaw, Russell J. Mumper, and A. T. Quantitative Nanostructure-Activity Relationship (QNAR) Modeling. *ACS Nano* **5**, 1–7 (2010).
138. Toropova, A. P. *et al.* Optimal descriptor as a translator of eclectic information into the prediction of membrane damage: The case of a group of ZnO and TiO<sub>2</sub> nanoparticles. *Ecotoxicol. Environ. Saf.* **108**, 203–209 (2014).
139. To, K. T., Truong, L., Edwards, S., Tanguay, R. L. & Reif, D. M. Multivariate modeling of engineered nanomaterial features associated with developmental toxicity. *NanoImpact* **16**, (2019).
140. Roy, J., Ojha, P. K. & Roy, K. Risk assessment of heterogeneous TiO<sub>2</sub>-based engineered nanoparticles (NPs): a QSTR approach using simple periodic table based descriptors. *Nanotoxicology* **13**, 701–716 (2019).
141. González-Durruthy, M. *et al.* Experimental-Computational Study of Carbon Nanotube Effects on Mitochondrial Respiration: In Silico Nano-QSPR Machine Learning Models Based on New Raman Spectra Transform with Markov-Shannon Entropy Invariants. *J. Chem. Inf. Model.* **57**, 1029–1044

- (2017).
142. Mikolajczyk, A. *et al.* Zeta Potential for Metal Oxide Nanoparticles: A Predictive Model Developed by a Nano-Quantitative Structure–Property Relationship Approach. *Chem. Mater.* **27**, 2400–2407 (2015).
  143. Caruana, R. & Niculescu-Mizil, A. Data mining in metric space. in *Proceedings of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '04* 69 (ACM Press, 2004). doi:10.1145/1014052.1014063.
  144. Roy, K., Das, R. N., Ambure, P. & Aher, R. B. Be aware of error measures. Further studies on validation of predictive QSAR models. *Chemom. Intell. Lab. Syst.* **152**, 18–33 (2016).
  145. Hand, D. & Christen, P. A note on using the F-measure for evaluating record linkage algorithms. *Stat. Comput.* **28**, 539–547 (2018).
  146. Chicco, D. & Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics* **21**, 1–13 (2020).
  147. Jaworska, J., Nikolova-Jeliazkova, N. & Aldenberg, T. QSAR applicability domain estimation by projection of the training set in descriptor space: A review. *ATLA Altern. to Lab. Anim.* **33**, 445–459 (2005).
  148. Sahigara, F. *et al.* Comparison of different approaches to define the applicability domain of QSAR models. *Molecules* **17**, 4791–4810 (2012).
  149. Gadaleta, D., Mangiatordi, G. F., Catto, M., Carotti, A. & Nicolotti, O. Applicability Domain for QSAR Models: Where Theory Meets Reality. *Int. J. Quant. Struct. Relationships* **1**, 45–63 (2016).
  150. Gramatica, P. Principles of QSAR models validation: Internal and external. *QSAR Comb. Sci.* **26**, 694–701 (2007).
  151. Mikolajczyk, A. *et al.* Nano-QSAR modeling for ecosafe design of heterogeneous TiO<sub>2</sub>-based nano-photocatalysts. *Environ. Sci. Nano* **5**, 1150–1160 (2018).
  152. Gajewicz, A. Development of valuable predictive read-across models based on “real-life” (sparse) nanotoxicity data. *Environ. Sci. Nano* **4**, 1389–1403 (2017).
  153. Liu, R., Hassan, T., Rallo, R. & Cohen, Y. HDAT: web-based high-throughput screening data analysis tools. *Comput. Sci. Discov.* **6**, 014006 (2013).
  154. Papa, E., Doucet, J. P. & Doucet-Panaye, A. Linear and non-linear modelling of the cytotoxicity of TiO<sub>2</sub> and ZnO nanoparticles by empirical descriptors. *SAR QSAR Environ. Res.* **26**, 647–665 (2015).
  155. Papadiamantis, A. G. *et al.* Predicting cytotoxicity of metal oxide nanoparticles using isalos analytics platform. *Nanomaterials* **10**, 2017 (2020).
  156. Liu, R., Zhang, Y. H., Xia Ji, Z. & Rallo, R. Development of Structure-Activity Relationship for Metal Oxide Nanoparticles. *RSC Publ.* (2013).
  157. Gajewicz, A. *et al.* Decision tree models to classify nanomaterials according to the DF4nanoGrouping scheme. *Nanotoxicology* **0**, 1–17 (2017).
  158. Roy, K., Kar, S. & Ambure, P. On a simple approach for determining applicability domain of QSAR models. *Chemom. Intell. Lab. Syst.* **145**, 22–29 (2015).
  159. Chau, Y. T. & Yap, C. W. Quantitative Nanostructure-Activity Relationship modelling of nanoparticles. *RSC Adv.* **2**, 8489–8496 (2012).

160. The European Materials Modelling Council. Data Documentation. <https://emmc.eu/moda/>.
161. Tsiros, P. *et al.* Towards an in silico integrated approach for testing and assessment of nanomaterials: from predicted indoor air concentrations to lung dose and biodistribution. *Environ. Sci. Nano* (2022) doi:10.1039/d1en00956g.
162. Cherkasov, A. *et al.* QSAR modeling: Where have you been? Where are you going to? *J. Med. Chem.* **57**, 4977–5010 (2014).
163. Worth, A. P. The role of QSAR methodology in the regulatory assessment of chemicals. in *Recent Advances in QSAR Studies* vol. 8 367–382 (2010).
164. Triebe, J., Janusch-Roi, A., Coe, A. & Worth, A. *JRC QSAR Model Database*. <https://publications.jrc.ec.europa.eu/repository/bitstream/JRC107491/kjna28713enn.pdf> (2017) doi:10.2760/905519.
165. Coady, K. *et al.* When Are Adverse Outcome Pathways and Associated Assays “Fit for Purpose” for Regulatory Decision-Making and Management of Chemicals? *Integr. Environ. Assess. Manag.* **15**, 633–647 (2019).
166. Serra, A. *et al.* Transcriptomics in Toxicogenomics, Part III: Data Modelling for Risk Assessment. *Nanomaterials* **10**, 708 (2020).
167. Serra, A. *et al.* INSIDE NANO: a systems biology framework to contextualize the mechanism-of-action of engineered nanomaterials. *Sci. Rep.* **9**, 179 (2019).
168. Li, D. *et al.* Physiologically based pharmacokinetic modeling of polyethylene glycol-coated polyacrylamide nanoparticles in rats. *Nanotoxicology* **8**, 128–137 (2014).
169. Li, M., Panagi, Z., Avgoustakis, K. & Reineke, J. Physiologically based pharmacokinetic modeling of PLGA nanoparticles with varied mPEG content. *Int. J. Nanomedicine* **7**, 1345–1356 (2012).
170. Lin, Z., Monteiro-Riviere, N. A., Kannan, R. & Riviere, J. E. A computational framework for interspecies pharmacokinetics, exposure and toxicity assessment of gold nanoparticles. *Nanomedicine* **11**, 107–119 (2016).
171. Zazo, H. *et al.* Physiologically Based Pharmacokinetic (PBPK) Model of Gold Nanoparticle-Based Drug Delivery System for Stavudine Biodistribution. *Pharmaceutics* **14**, (2022).
172. Péry, A. R. R., Brochot, C., Hoet, P. H. M., Nemmar, A. & Bois, F. Y. Development of a physiologically based kinetic model for 99m-Techetium-labelled carbon nanoparticles inhaled by humans Human PBPK model for carbon nanoparticles. *Inhal. Toxicol.* **21**, 1099–1107 (2009).
173. Chen, W. Y. *et al.* Physiologically based pharmacokinetic modeling of zinc oxide nanoparticles and zinc nitrate in mice. *Int. J. Nanomedicine* **10**, 6277–6292 (2015).
174. Sahneh, F. D., Scoglio, C. M., Monteiro-Riviere, N. A. & Riviere, J. E. Predicting the impact of biocorona formation kinetics on interspecies extrapolations of nanoparticle biodistribution modeling. *Nanomedicine* **10**, 25–33 (2015).
175. Tu, Y. *et al.* Destructive extraction of phospholipids from Escherichia coli membranes by graphene nanosheets. *Nat. Nanotechnol.* (2013) doi:10.1038/nnano.2013.125.
176. Duan, G. *et al.* Protein corona mitigates the cytotoxicity of graphene oxide by reducing its physical interaction with cell membrane. *Nanoscale* **7**, 15214–15224 (2015).
177. Li, Y. *et al.* Graphene microsheets enter cells through spontaneous membrane penetration at edge asperities and corner sites. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 12295–12300 (2013).

178. Zhu, X., Li, N., Huang, C., Li, Z. & Fan, J. Membrane Perturbation and Lipid Flip-Flop Mediated by Graphene Nanosheet. *J. Phys. Chem. B* (2020) doi:10.1021/acs.jpcb.0c06089.
179. Dallavalle, M., Bottoni, A., Calvaresi, M. & Zerbetto, F. Functionalization Pattern of Graphene Oxide Sheets Controls Entry or Produces Lipid Turmoil in Phospholipid Membranes. *ACS Appl. Mater. Interfaces* **10**, 15487–15493 (2018).
180. Li, Z., Zhang, Y., Ma, J., Meng, Q. & Fan, J. Modeling Interactions between Liposomes and Hydrophobic Nanosheets. *Small* **15**, 1–10 (2019).
181. Puigpelat, E., Ignés-Mullol, J., Sagués, F. & Reigada, R. Interaction of Graphene Nanoparticles and Lipid Membranes Displaying Different Liquid Orderings: A Molecular Dynamics Study. *Langmuir* **35**, 16661–16668 (2019).
182. Chen, J. *et al.* Interaction of Graphene and its Oxide with Lipid Membrane: A Molecular Dynamics Simulation Study. *J. Phys. Chem. C* (2016) doi:10.1021/acs.jpcc.5b10635.
183. Chong, Y. *et al.* Reduced Cytotoxicity of Graphene Nanosheets Mediated by Blood-Protein Coating. *ACS Nano* **9**, 5713–5724 (2015).
184. Feng, M., Kang, H., Yang, Z., Luan, B. & Zhou, R. Potential disruption of protein-protein interactions by graphene oxide. *J. Chem. Phys.* **144**, (2016).
185. Zhu, W. *et al.* Nanomechanical mechanism for lipid bilayer damage induced by carbon nanotubes confined in intracellular vesicles. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 12374–12379 (2016).
186. Melby, E. S. *et al.* Peripheral Membrane Proteins Facilitate Nanoparticle Binding at Lipid Bilayer Interfaces. *Langmuir* **34**, 10793–10805 (2018).
187. Van Lehn, R. C. & Alexander-Katz, A. Pathway for insertion of amphiphilic nanoparticles into defect-free lipid bilayers from atomistic molecular dynamics simulations. *Soft Matter* **11**, 3165–3175 (2015).
188. Halappanavar, S. *et al.* Adverse outcome pathways as a tool for the design of testing strategies to support the safety assessment of emerging advanced materials at the nanoscale. *Part. Fibre Toxicol.* **17**, 16 (2020).
189. Labib, S. *et al.* Nano-risk Science: Application of toxicogenomics in an adverse outcome pathway framework for risk assessment of multi-walled carbon nanotubes. *Part. Fibre Toxicol.* **13**, 1–17 (2016).
190. Vietti, G., Lison, D. & van den Brule, S. Mechanisms of lung fibrosis induced by carbon nanotubes: Towards an Adverse Outcome Pathway (AOP). *Part. Fibre Toxicol.* **13**, (2016).
191. Grafström, R. C. *et al.* Toward the replacement of animal experiments through the bioinformatics-driven analysis of ‘omics’ data from human cell cultures. *ATLA Altern. to Lab. Anim.* **43**, 325–332 (2015).
192. Nymark, P. *et al.* A Data Fusion Pipeline for Generating and Enriching Adverse Outcome Pathway Descriptions. *Toxicol. Sci.* **162**, 264–275 (2018).
193. Nymark, P., Kohonen, P., Hongisto, V. & Grafström, R. C. Toxic and genomic influences of inhaled nanomaterials as a basis for predicting adverse outcome. in *Annals of the American Thoracic Society* vol. 15 S91–S97 (American Thoracic Society, 2018).
194. Halappanavar, S. *et al.* Ranking of nanomaterial potency to induce pathway perturbations associated with lung responses. *NanoImpact* **14**, 100158 (2019).

195. Liu, H. H., Bilal, M., Lazareva, A., Keller, A. & Cohen, Y. Simulation tool for assessing the release and environmental distribution of nanomaterials. *Beilstein J. Nanotechnol.* **6**, 938–951 (2015).
196. Zheng, Y. & Nowack, B. Size-specific, dynamic, probabilistic material flow analysis of titanium dioxide releases into the environment. *Environ. Sci. Technol.* **55**, 2392–2402 (2021).
197. Liu, R. *et al.* Evaluation of Toxicity Ranking for Metal Oxide Nanoparticles via an in Vitro Dosimetry Model. *ACS Nano* **9**, 9303–9313 (2015).
198. NanoSolveIT. NanoSolveIT Cloud Platform. <https://cloud.nanosolveit.eu/>.
199. Ambure, P., Balasaheb, R., Gajewicz, A., Puzyn, T. & Roy, K. ‘NanoBRIDGES’ software: Open access tools to perform QSAR and nano-QSAR modeling. *Chemom. Intell. Lab. Syst.* **147**, 1–13 (2015).
200. Mollá, F. A., Fito-López, C., Alvaro, J. A. H. & Huertas-López, F. New tools to support the risk assessment process of nanomaterials in the insurance sector. *Int. J. Environ. Res. Public Health* **18**, (2021).
201. Shin, H. K., Kim, S. & Yoon, S. Use of size-dependent electron configuration fingerprint to develop general prediction models for nanomaterials. *NanoImpact* **21**, 100298 (2021).
202. Ruusmann, V., Sild, S. & Maran, U. QSAR DataBank repository: Open and linked qualitative and quantitative structure-activity relationship models. *J. Cheminform.* **7**, 1–11 (2015).
203. Cheimarios, N. *et al.* An In Vitro Dosimetry Tool for the Numerical Transport Modeling of Engineered Nanomaterials Powered by the Enalos RiskGONE Cloud Platform. *Nanomaterials* **12**, 3935 (2022).
204. Roy, K., Ambure, P., Kar, S. & Ojha, P. K. Is it possible to improve the quality of predictions from an “intelligent” use of multiple QSAR/QSPR/QSTR models? *J. Chemom.* **32**, e2992 (2018).
205. Fourches, D., Muratov, E. & Tropsha, A. Trust, but Verify II: A Practical Guide to Chemogenomics Data Curation. *J. Chem. Inf. Model.* **56**, 1243–1252 (2016).
206. Vijayabaskar, M. S. & Vishveshwara, S. Interaction energy based protein structure networks. *Biophys. J.* **99**, 3704–3715 (2010).
207. Daszykowski, M., Walczak, B. & Massart, D. L. Representative subset selection. *Anal. Chim. Acta* (2002) doi:10.1016/S0003-2670(02)00651-7.
208. Golbraikh, A. & Tropsha, A. Beware of  $q^2$ ! *J. Mol. Graph. Model.* **20**, 269–276 (2002).
209. Todeschini, R. & Consonni, V. *Molecular Descriptors for Chemoinformatics*. vol. 2 (Wiley, 2009).
210. Sundén, B. Hydrogen. in *Hydrogen, Batteries and Fuel Cells* 37–55 (Elsevier, 2019). doi:10.1016/B978-0-12-816950-6.00003-8.
211. Niaz, S., Manzoor, T. & Pandith, A. H. Hydrogen storage: Materials, methods and perspectives. *Renew. Sustain. Energy Rev.* **50**, 457–469 (2015).
212. Allied Market Research. Nanoparticle Titanium Dioxide Market Outlook - 2021-2030. <https://www.alliedmarketresearch.com/nanoparticle-titanium-dioxide-market-A11862>.
213. Shi, H., Magaye, R., Castranova, V. & Zhao, J. Titanium dioxide nanoparticles: a review of current toxicological data. *Part. Fibre Toxicol.* **10**, 15 (2013).
214. Waghmode, M. S., Gunjal, A. B., Mulla, J. A., Patil, N. N. & Nawani, N. N. Studies on the titanium dioxide nanoparticles: biosynthesis, applications and remediation. *SN Appl. Sci.* **1**, 1–9 (2019).



215. Ziental, D. *et al.* Titanium dioxide nanoparticles: Prospects and applications in medicine. *Nanomaterials* **10**, (2020).
216. Janczarek, M. *et al.* Progress of functionalized TiO<sub>2</sub>-based nanomaterials in the construction industry: A comprehensive review. *Chem. Eng. J.* **430**, (2022).
217. Toropov, A., Sizochenko, N., Toropova, A. & Leszczynski, J. Towards the Development of Global Nano-Quantitative Structure–Property Relationship Models: Zeta Potentials of Metal Oxide Nanoparticles. *Nanomaterials* **8**, 243 (2018).
218. Organisation for Economic Co-operation and Development. OECD Principles for The Validation, For Regulatory Purposes, of (Quantitative) Structure-Activity Relationship Models. 3–4 (2004).
219. Joossens, E. *et al.* A high throughput imaging database of toxicological effects of nanomaterials tested on HepaRG cells. *Sci. Data* **6**, 46 (2019).
220. Pathakoti, K., Huang, M.-J., Watts, J. D., He, X. & Hwang, H.-M. Using experimental data of Escherichia coli to develop a QSAR model for predicting the photo-induced cytotoxicity of metal oxide nanoparticles. *J. Photochem. Photobiol. B Biol.* **130**, 234–240 (2014).
221. Falk, A. *et al.* ‘Closer to the Market’ Roadmap (CTTM). *NanoSafety Clust.* 1–43 (2016) doi:10.5281/ZENODO.1493492.
222. Chen, R. *et al.* Nanoparticle surface characterization and clustering through concentration-dependent surface adsorption modeling. *ACS Nano* **8**, 9446–9456 (2014).
223. Fourches, D. *et al.* Quantitative Nanostructure–Activity Relationship Modeling. *ACS Nano* **4**, 5703–5712 (2010).
224. Sayes, C. & Ivanov, I. Comparative Study of Predictive Computational Models for Nanoparticle-Induced Cytotoxicity. *Risk Anal.* **30**, 1723–1734 (2010).
225. Liu, X. *et al.* Predictive modeling of nanomaterial exposure effects in biological systems. *Int. J. Nanomedicine* **8**, 31–43 (2013).
226. Borders, T. L., Fonseca, A. F., Zhang, H., Cho, K. & Rusinko, A. Developing descriptors to predict mechanical properties of nanotubes. *J. Chem. Inf. Model.* **53**, 773–782 (2013).
227. Guo, H. *et al.* An ensemble learning for predicting breakdown field strength of polyimide nanocomposite films. *J. Nanomater.* **2015**, (2015).
228. Fazeli Sangani, M. *et al.* Different modelling approaches for predicting titanium dioxide nanoparticles mobility in intact soil media. *Sci. Total Environ.* **665**, 1168–1181 (2019).
229. Timoshenko, J., Lu, D., Lin, Y. & Frenkel, A. I. Supervised Machine-Learning-Based Determination of Three-Dimensional Structure of Metallic Nanoparticles. *J. Phys. Chem. Lett.* **8**, 5091–5098 (2017).
230. Marvin, H. J. P. *et al.* Application of Bayesian networks for hazard ranking of nanomaterials to support human health risk assessment. *Nanotoxicology* **11**, 123–133 (2017).
231. Gasper, R., Shi, H. & Ramasubramaniam, A. Adsorption of CO on Low-Energy, Low-Symmetry Pt Nanoparticles: Energy Decomposition Analysis and Prediction via Machine-Learning Models. *J. Phys. Chem. C* **121**, 5612–5619 (2017).
232. Speck-Planche, A., Kleandrova, V. V., Luan, F. & Ds Cordeiro, M. N. Computational modeling in nanomedicine: Prediction of multiple antibacterial profiles of nanoparticles using a quantitative structure-activity relationship perturbation model. *Nanomedicine* **10**, 193–204 (2015).

233. Kleandrova, V. V. *et al.* Computational ecotoxicology: Simultaneous prediction of ecotoxic effects of nanoparticles under different experimental conditions. *Environ. Int.* **73**, 288–294 (2014).
234. Brinkmann, B. W. *et al.* Predicted Adsorption Affinity for Enteric Microbial Metabolites to Metal and Carbon Nanomaterials. *J. Chem. Inf. Model.* **62**, 3589–3603 (2022).
235. Fanourgakis, G. S., Gkagkas, K., Tylanakis, E. & Froudakis, G. A Generic Machine Learning Algorithm for the Prediction of Gas Adsorption in Nanoporous Materials. *J. Phys. Chem. C* **124**, 7117–7126 (2020).
236. Roy, J. & Roy, K. Nano-read-across predictions of toxicity of metal oxide engineered nanoparticles (MeOx ENPS) used in nanopesticides to BEAS-2B and RAW 264.7 cells. *Nanotoxicology* **16**, 629–644 (2022).
237. Toropov, A. A. *et al.* QSAR as a random event: Modeling of nanoparticles uptake in PaCa2 cancer cells. *Chemosphere* **92**, 31–37 (2013).
238. Toropova, A. P., Toropov, A. A., Benfenati, E. & Korenstein, R. QSAR model for cytotoxicity of SiO<sub>2</sub> nanoparticles on human lung fibroblasts. *J. Nanoparticle Res.* **16**, (2014).
239. Toropov, A. A. & Toropova, A. P. Quasi-SMILES and nano-QFAR: United model for mutagenicity of fullerene and MWCNT under different conditions. *Chemosphere* **139**, 18–22 (2015).
240. Baharifar, H. & Amani, A. Cytotoxicity of chitosan/streptokinase nanoparticles as a function of size: An artificial neural networks study. *Nanomedicine Nanotechnology, Biol. Med.* **12**, 171–180 (2016).
241. Sattari Baboukani, B., Ye, Z., G. Reyes, K. & Nalam, P. C. Prediction of Nanoscale Friction for Two-Dimensional Materials Using a Machine Learning Approach. *Tribol. Lett.* **68**, 1–14 (2020).
242. Bilal, M. *et al.* Bayesian Network Resource for Meta-Analysis: Cellular Toxicity of Quantum Dots. *Small* **15**, 1–12 (2019).
243. Oh, E. *et al.* Meta-analysis of cellular toxicity for cadmium-containing quantum dots. *Nat. Nanotechnol.* **11**, 479–486 (2016).
244. Gernand, J. M. & Casman, E. A. A meta-analysis of carbon nanotube pulmonary toxicity studies-how physical dimensions and impurities affect the toxicity of carbon nanotubes. *Risk Anal.* **34**, 583–597 (2014).
245. Choi, J. S., Ha, M. K., Trinh, T. X., Yoon, T. H. & Byun, H. G. Towards a generalized toxicity prediction model for oxide nanomaterials using integrated data from different sources. *Sci. Rep.* **8**, 1–10 (2018).
246. Fan, W., Liu, L., Peng, R. & Wang, W.-X. High bioconcentration of titanium dioxide nanoparticles in *Daphnia magna* determined by kinetic approach. *Sci. Total Environ.* **569–570**, 1224–1231 (2016).
247. Garner, K. L., Qin, Y., Cucurachi, S., Suh, S. & Keller, A. A. Linking Exposure and Kinetic Bioaccumulation Models for Metallic Engineered Nanomaterials in Freshwater Ecosystems. *ACS Sustain. Chem. Eng.* **6**, 12684–12694 (2018).
248. Kreyling, W. G. *et al.* Quantitative biokinetics over a 28 day period of freshly generated, pristine, 20 nm silver nanoparticle aerosols in healthy adult rats after a single 1½-hour inhalation exposure. *Part. Fibre Toxicol.* **17**, 21 (2020).
249. DeLoid, G. M. *et al.* Advanced computational modeling for in vitro nanomaterial dosimetry. *Part.*

- Fibre Toxicol.* **12**, 1–20 (2015).
250. Deloid, G. M., Cohen, J. M., Pyrgiotakis, G. & Demokritou, P. Preparation, characterization, and in vitro dosimetry of dispersed, engineered nanomaterials. *Nat. Protoc.* **12**, 355–371 (2017).
  251. Jensen, A. C. Ø. *et al.* Comparison of geometrical layouts for a multi-box aerosol model from a single-chamber dispersion study. *Environ. - MDPI* **5**, 1–16 (2018).
  252. Golbraikh, A. & Tropsha, A. Beware of the q<sup>2</sup>! *J. Mol. Graph. Model.* **20**, 269–276 (2002).
  253. Organisation for Economic Co-operation and Development. Moving Towards a Safe(r) Innovation Approach (SIA) for More Sustainable Nanomaterials and Nano-enabled Products. Series on the Safety of Manufactured Nanomaterials. No. 96. (2020).
  254. Park, C.-B. *et al.* Mixture toxicity of metal oxide nanoparticles and silver ions on *Daphnia magna*. *J. Nanoparticle Res.* **21**, 166 (2019).

## Appendix I – List of assessed studies and models for nanoQSAR, grouping and read-across alternative methods

Table 21: List of assessed studies and models for nanoQSAR, grouping and read-across alternative methods. Some key information is included in this Table. The “Resource ID” column refers to the models ID as recorded in the master file and next to the quality assessment task. In the “Endpoint” column [R] refers to regression (continuous/numerical endpoint) and [C] refers to classification (categorical endpoint). In each study more than one dataset may have been employed. The “Model documentation” column refers to the availability of a QMRF or MODA report for the respective model.

Resource ID	Authors	Endpoint	Dataset size	Variable selection	Modelling approach	Model validation	Statistics (at least 4)	DoA	Robustness and/or variability assessment	Model interpretation	Model documentation	Data availability
1	Varsou <i>et al.</i> <sup>37</sup>	[R] Cell association in A549 cell line (log <sub>2</sub> -transformed) - [General methodology that can be used with different data and produce different models]	84	Yes	The prediction for an untested NM is the weighted average of similar NMs	Internal validation	No	No	Yes	No	No	Yes
2	Varsou <i>et al.</i> <sup>89</sup>	[R] Zeta potential in water (pH=6.5 or 7)	37	Yes	kNN	External validation	Yes	Yes	Yes	Yes	No	No
3	Varsou <i>et al.</i> <sup>34,123</sup>	[R] Cell association in A549 cell line (log <sub>2</sub> -transformed), [C] Toxicity profile, Cell viability [General methodology that can be used with different data and produce different models]	84, 24, 16	Yes	The prediction for an untested NM is the weighted average of similar NMs (numerical endpoint) or the weighted majority vote of similar NMs	Internal and external validation	Yes	Yes	Yes	No	No	Yes

4	Papadiamantis <i>et al.</i> <sup>1</sup>	[R] Zeta potential in water (pH=7)	69	Yes	kNN	Internal and external validation	Yes	Yes	Yes	Yes	Yes	Yes
5	Varsou <i>et al.</i> <sup>35</sup>	[R] Cell association in A549 cell line (log <sub>2</sub> -transformed), adsorption coefficient (log-transformed)- [General methodology that can be used with different data and produce different models]	84, 28	Yes	In each group the prediction is calculated by the corresponding MLR model	Internal and external validation	Yes	Yes	Yes	No	No	Yes
6	Kotzabasa <i>et al.</i> <sup>76</sup>	[C] Cell viability of STEM cells	16	Yes	Logistic regression model	Internal and external validation	Yes	Yes	Yes	No	Yes	Yes
7	Varsou <i>et al.</i> <sup>70</sup>	[C] Protein binding to carbonic anhydrase (CA), cell viability by WST-1 assay	83	Yes	kNN	Internal and external validation	Yes	Yes	Yes	Yes	No	Yes
8	Varsou <i>et al.</i> <sup>75</sup>	[C] Cytotoxicity to <i>Daphnia Magna</i> (death at 48hrs) expressed as the EC <sub>40</sub> concentration	353	Yes	kNN	Internal and external validation	Yes	Yes	Yes	Yes	Yes	Yes
9	Melagraki <i>et al.</i> <sup>65</sup>	[C] Toxicity class ("active"/ "inactive") summarized after considering 64 bioactivity measures	44	No	J48 classification tree	Internal and external validation	Yes	Yes	Yes	No	No	Yes
10	Kotzabasa <i>et al.</i> <sup>77</sup>	[C] Genotoxicity class	15	Yes	Logistic regression model	Internal and external validation	Yes	Yes	Yes	No	Yes	Yes
11	Afantitis <i>et al.</i> <sup>36</sup>	[R] Cell association in A549 cell line (log <sub>2</sub> -transformed)	105	Yes	kNN	Internal and external validation	Yes	Yes	Yes	Yes	No	Yes
12	Melagraki <i>et al.</i> <sup>44</sup>	[R] Cellular uptake in pancreatic cancer	109	Yes	kNN	Internal and external validation	Yes	Yes	Yes	No	No	No

		cells (PaCa2) (log-transformed)										
13	Puzyn <i>et al.</i> <sup>50</sup>	[R] Cytotoxicity to <i>E. Coli</i> (logarithmic values of molar 1/EC <sub>50</sub> )	17	Yes	MLR combined with a genetic algorithm	Internal and external validation	Yes	Yes	Yes	Yes	No	Yes
14	Papadimitriou <i>et al.</i> <sup>155</sup>	[R] Cytotoxicity (% cell viability) to human bronchial epithelial (BEAS-2B) and murine myeloid (RAW264.7) cell lines	345	Yes	kNN	External validation	Yes	Yes	Yes	Yes	Yes	Yes
15	Kar <i>et al.</i> <sup>45</sup>	[R] Cellular uptake in pancreatic cancer cells (PaCa2) (log-transformed)	109	Yes	Stepwise MLR followed by PLS regression	Internal and external validation	Yes	Yes	Yes	Yes	No	No
16	Kar <i>et al.</i> <sup>52</sup>	[C] Cytotoxicity to <i>E. Coli</i> ("low-L"/"high-H")	25	Yes	Linear discriminant analysis (LDA) - best performing method	External validation	Yes	Yes	No	Yes	No	Yes
17	Papa <i>et al.</i> <sup>41</sup>	[R] Cell association in A549 cell line (log <sub>2</sub> -transformed)	84	Yes	MLR, SVM, PLS, projection pursuit regression (PPR), kNN, multivariate adaptive regression splines (MARS), Neural Networks (RBFNN, GRegNN, CPANN),	Internal and external validation	Yes	Yes	Yes	Yes	No	Yes

					random forest. PPR presents the best performance.							
18	Sang <i>et al.</i> <sup>90</sup>	[R] Cell viability of human renal cortex proximal tubule epithelial (HK-2) cells	72	Yes	Random forest and K-means clustering analysis (best performing method)	Internal and external validation	Yes	Yes	Yes	Yes	No	Yes
19	Thwala <i>et al.</i> <sup>88</sup>	[C] Toxicity to human BEAS-2B, murine myeloid (RAW 264.7) cell lines, and <i>E. Coli</i>	26	Yes	kNN	External validation	Yes	Yes	Yes	Yes	No	Yes
20	Swirog <i>et al.</i> <sup>103</sup>	[R] Electrophoretic mobility ( $\mu\text{m s}^{-1}$ V <sup>-1</sup> cm)	27	Yes	GA-PLS (best performing method)	Internal and external validation	Yes	Yes	Yes	Yes	No	Yes
21	Cai <i>et al.</i> <sup>91</sup>	[R] Toxicity to <i>Daphnia Magna</i> expressed as EC <sub>50</sub> concentration	23	Yes	MLR	Training and validation using the whole set	No	No	Yes	Yes	No	Yes
22	Qi <i>et al.</i> <sup>92</sup>	[R] Cytotoxicity to Chinese hamster ovary cells (logarithmic values of molar 1/EC <sub>50</sub> )	34	No	Monte Carlo-PLS	Internal and external validation	Yes	Yes	Yes	Yes	No	No

23	Balraadjising <i>et al.</i> <sup>118</sup>	[C] Acute immobilization ( <i>in vivo</i> ) of <i>Daphnia Magna</i> expressed in EC <sub>50</sub> values and transformed to classes ("very toxic", "toxic", "harmful", "not harmful") after exposure to metallic NMs	454	Yes	Random forests (best performing method)	Internal and external validation	Yes	Yes	Yes	Yes	No	No
24	Chatterjee <i>et al.</i> <sup>59</sup>	[R] Cytotoxicity to human keratinocyte cell line (HaCaT) expressed as the molar 1/log LC <sub>50</sub> , cytotoxicity to <i>E. Coli</i> (logarithmic values of molar 1/EC <sub>50</sub> or 1/log LC <sub>50</sub> )- [General methodology that can be used with different data and produce different models]	18,17,16	Yes	The prediction for an untested NM is the weighted average of similar NMs	External validation	Yes	No	No	No	No	Yes
25	Forest <i>et al.</i> <sup>104</sup>	[C/R] Cytotoxicity ("toxic"/"nontoxic") based on the assessment of the LDH release from RAW 264.7 cells	25	Yes	PLS regression and CART (Classification and Regression Tree)	Internal validation	Yes	Yes	Yes	Yes	No	Yes
26	Liu <i>et al.</i> <sup>110</sup>	[C] Cytotoxicity via propidium iodide uptake of BEAS-2B cells	9 NMs	Yes	Logistic regression model	Internal and external validation	No	Yes	Yes	Yes	No	Yes
27	Pathakoti <i>et al.</i> <sup>220</sup>	[R] Cytotoxicity to <i>E. Coli</i> (logarithmic values of molar 1/LC <sub>50</sub> )	17	Yes	Least square regression	Internal and external validation	No	No	Yes	Yes	No	Yes



28	George <i>et al.</i> <sup>93</sup>	[C] <i>In vitro</i> sublethal and lethal endpoints and <i>in vivo</i> response on Zebrafish embryo	4032	No	SOM	None reported	No	No	No	Yes	No	No
29	Zhang <i>et al.</i> <sup>78</sup>	[R] <i>In vitro</i> oxidative stress, acute pulmonary inflammation	24	Yes	Regression tree	Internal validation	No	No	Yes	Yes	No	No
30	Chen <i>et al.</i> <sup>222</sup>	[R] Biological surface adsorption index (BSAI) denoted as logk values of adsorption coefficient in organic molecules	23	Yes	Linear free energy relationship and PCA	Internal and external validation	No	Yes	Yes	Yes	No	Yes
32	Gajewicz <i>et al.</i> <sup>26</sup>	[C] Cytotoxicity to <i>E. Coli</i> (logarithmic values of molar 1/EC50), Cytotoxicity to human keratinocyte cell line (HaCaT) expressed as the molar 1/log LC50. Values were transformed to classes. - [General methodology that can be used with different data and produce different models]	17, 18	Yes	Hierarchical clustering	External validation	No	No	No	Yes	No	Yes
33	Gajewicz <i>et al.</i> <sup>57</sup>	[R] Cytotoxicity to <i>E. Coli</i> (logarithmic values of molar 1/EC <sub>50</sub> ), Cytotoxicity to human keratinocyte cell line (HaCaT) expressed as the molar 1/log LC <sub>50</sub> - [General methodology that can be used with different	17, 18	Yes	One-point slope, two points formula, plane passing for three points depending on the number of	External validation	Yes	No	No	No	No	Yes

		data and produce different models]			the available descriptors							
34	Helma <i>et al.</i> <sup>38</sup>	[R] Cell association in A549 cell line (log <sub>2</sub> -transformed)	121	Yes	Weighted local average, weighted PLS regression and weighted random forests	Internal validation	No	Yes	Yes	No	No	No
35	Sizochenko <i>et al.</i> <sup>56</sup>	[R] <i>In vitro</i> cytotoxicity data for bacteria, algae, protozoa, human keratinocyte cell, Balb/c 3T3, expressed as logarithmic values of molar 1/EC <sub>50</sub> or 1/IC <sub>50</sub>	30	No	Pearson correlation coefficient (similarity), SOM	None reported	No	No	No	Yes	No	Yes
36	Lamon <i>et al.</i> <sup>113</sup>	[C] <i>In vitro</i> genotoxicity via Comet assay	8	Yes	PCA, hierarchical clustering, random forest	None reported	No	Yes	Yes	Yes	No	Yes

37	Aschberger <i>et al.</i> <sup>102</sup>	[C] <i>In vivo</i> and <i>in vitro</i> genotoxicity	19	Yes	PCA, hierarchical clustering. Endpoint estimation: All assessed analogues form one category for this toxicological endpoint and data gaps could be filled by interpolation.	None reported	No	Yes	Yes	Yes	No	Yes
38	Epa <i>et al.</i> <sup>48</sup>	[R] Smooth muscle apoptosis (SMA), Cellular uptake in pancreatic cancer (PaCa2) and human umbilical vein endothelial cell (HUVEC)	31, 109	Yes	MLR, Bayesian-regularized ANN methods	External validation	Yes	No	No	Yes	No	No
39	Mikolajczyk <i>et al.</i> <sup>100</sup>	[R] The efficiency of phenol degradation [%] in UV-vis light, Cytotoxicity on epithelial cells from Chinese hamster ovary (CHO-K1 cell line, ATCC® CCL-61™). The final values of the cytotoxicity were expressed as the logarithm of 50% inhibition of the cell viability (pEC <sub>50</sub> ).	29	Yes	Linear MLR-GA and non-linear decision tree (DT)	Internal and external validation	Yes	Yes	Yes	Yes	No	Yes

40	Stolinski <i>et al.</i> <sup>101</sup>	[R] Cytotoxicity on epithelial cells from Chinese hamster ovary (CHO-K1 cell line, ATCC® CCL-61™). The final values of the cytotoxicity were expressed as the logarithm of 50% inhibition of the cell viability (pEC <sub>50</sub> ).	29	Yes	MLR	Performed on previous publication	No	Yes	No	No	No	Yes
41	Rybinska-Fryca <i>et al.</i> <sup>112</sup>	[C/R] Photodegradation of phenol by titanium dioxide (TiO <sub>2</sub> ). A threshold is used to define the classes of high or low photoactivity (percent of degradation less than 35%).	21	Yes	Logistic regression model, PCA	Internal and external validation	Yes	Yes	Yes	No	No	Yes
42	Wang <i>et al.</i> <sup>73</sup>	[R] Root concentration factor (RCF) and translocation factor (TF)	114, 88, 106, 130	Yes	Back propagation neural network (BPNN)	External validation (using 3 subsets)	No	No	Yes	Yes	No	Yes
43	Fourches <i>et al.</i> <sup>223</sup>	[C] Toxicity class ("active"/ "inactive") summarized after considering 64 bioactivity measures, [R] Cellular uptake in pancreatic cancer cells (PaCa2) (log-transformed)	44, 109	Yes	SVM classification, kNN-based regression, hierarchical clustering analysis (HCA)	Internal and External validation	Yes	Yes	Yes	Yes	No	Yes
44	Kar <i>et al.</i> <sup>51</sup>	[R] Cytotoxicity to <i>E. Coli</i> (logarithmic values of molar 1/EC <sub>50</sub> ), Cytotoxicity	17, 18	Yes	Stepwise MLR, PLS	Internal validation	Yes	Yes	Yes	Yes	No	Yes

		to human keratinocyte cell line (HaCaT) expressed as the molar 1/log LC <sub>50</sub>										
45	Papa <i>et al.</i> <sup>154</sup>	[C/R] Membrane damage in cells measured by the release of LDH	42	Yes	MLR based on ordinary least squares (MLR-OLS), SVM, Neural networks (RBFNN and GRNN), J48 classification tree	Internal and External validation	Yes	Yes	Yes	Yes	No	Yes
46	Walkey <i>et al.</i> <sup>33</sup>	[R] Cell association in A549 cell line (log2-transformed)	105	Yes	PLS regression	Internal and external validation	Yes	Yes	Yes	Yes	No	Yes
47	Gajewicz <i>et al.</i> <sup>157</sup>	[C] Ferric reducing ability of serum (IOP), protein carbonylation, NOAEC from a short-term inhalation study in rats. Each endpoint is transformed in a binary class (active/inactive) according to reported thresholds	19	Yes	Decision tree models	External validation	Yes	Yes	Yes	Yes	No	No
48	Gajewicz <i>et al.</i> <sup>63</sup>	[R] Cytotoxicity to human keratinocyte cell line (HaCaT) expressed as the molar 1/log LC <sub>50</sub>	18	Yes	MLR combined with a genetic algorithm (GA-MLR)	Internal and external validation	Yes	Yes	Yes	Yes	No	Yes
49	Burello and Worth <sup>94</sup>	[R] Oxidative stress potential	70	No	NA	None reported	No	No	No	Yes	No	No

50	Sayes and Ivanov <sup>224</sup>	[C/R] Cellular membrane damage via LDH release	24, 18	Yes	MLR and linear discriminant analysis (LDA) based classification	Internal validation	No	No	Yes	Yes	No	Yes
51	Toropov <i>et al.</i> <sup>58</sup>	[R] Cytotoxicity to <i>E. Coli</i> (logarithmic values of molar 1/EC <sub>50</sub> )	17	Yes	MLR-Monte Carlo Optimization	Internal validation	No	No	Yes	No	No	No
52	Chau and Yap <sup>159</sup>	[C] Cellular uptake in pancreatic cancer (PaCa2)	105	Yes	Consensus classifier, Naive Bayes, Logistic Regression, kNN and SVM	Internal validation	No	No	Yes	No	No	Yes
53	Kleandrov <i>a et al.</i> <sup>121</sup>	[C] Toxic effect class (negative or positive) based on cut-off values for different measures of toxicity depending on biological target, namely Cytotoxicity, EC <sub>50</sub> , IC <sub>50</sub> , TC <sub>50</sub> , and LC <sub>50</sub> .	229	Yes	LDA	External validation	Yes	No	Yes	Yes	No	Yes
54	Kaweeteerawat <i>et al.</i> <sup>108</sup>	[C] Cytotoxicity to <i>E. Coli</i> viability by measuring growth inhibition (IC <sub>50</sub> ), membrane damage, and ROS generation	24	Yes	SVM classifier	None reported	No	No	Yes	Yes	No	No
55	Bigdeli <i>et al.</i> <sup>79</sup>	[R] Exocytosis in macrophages	12	Yes	PLS regression	Internal validation	No	No	Yes	Yes	No	Yes

56	Tsiliki <i>et al.</i> <sup>39</sup>	[R] Cell association in A549 cell line (log2-transformed), Cytotoxicity to <i>E. Coli</i> (logarithmic values of molar 1/EC50), Cytotoxicity to human keratinocyte cell line (HaCaT) expressed as the molar 1/log LC <sub>50</sub> - [General methodology that can be used with different data and produce different models]	84, 17, 18	Yes	SVM regression, Elastic Net Regression	Internal and external validation	No	Yes	Yes	No	No	Yes
57	Oksel <i>et al.</i> <sup>49</sup>	[C] Cell viability in BEAS-2B and RAW264.7 cell lines, cellular uptake in PaCa2 cancer cell line, cytotoxicity to via cellular uptake of HaCaT cells, exocytosis in macrophages- [General methodology that can be used with different data and produce different models]	24, 105, 18, 12	Yes	Decision Trees	External validation	Yes	No	Yes	Yes	No	Yes
58	Mu <i>et al.</i> <sup>60</sup>	[R] Cytotoxicity to <i>E. Coli</i> (logarithmic values of molar 1/EC <sub>50</sub> )	16+51 (untested)	Yes	MLR, Pearson's similarity, PCA and clustering analysis for variable selection	Internal and external validation	Yes	Yes	Yes	Yes	No	Yes

59	Pan <i>et al.</i> <sup>54</sup>	[R] Cytotoxicity to <i>E. Coli</i> (logarithmic values of molar 1/EC <sub>50</sub> ), Cytotoxicity to human keratinocyte cell line (HaCaT) expressed as the molar 1/log LC <sub>50</sub>	17, 18	No	Monte Carlo optimization	Internal and external validation	Yes	No	Yes	Yes	No	Yes
60	Serra <i>et al.</i> <sup>167</sup>	[Mode of action contextualization] Assessment of the connections between phenotypic entities based on their effects on the genes.	28 ENMs, 615 drugs, 585 human diseases and 2288 chemicals	No	Similarity search (Jaccard index), Gene set enrichment analysis, phenotypic network inference, cliques search	Comparison with other independently computed similarities based on different characteristics, such as the molecular structure of the drugs and chemicals, the symptoms of the diseases, the use in clinical practice of drugs, and the pathogenic roles of chemical exposures.	No	No	No	No	No	Yes



61	Gousiadou <i>et al.</i> <sup>117</sup>	[R] Toxicological responses observed in zebrafish ( <i>Danio Rerio</i> ) expressed as the Weighted and Additive EZ Metric based on the aggregation of a variety of biological endpoints	176 (44 NMs), 47 (12 NMs)	Yes	Random Forest, Decision Trees, Gradient Boosting Machine, kNN, Linear Regression, Generalized Linear Regression, SVM with Radial Function, XGB	Internal and external validation	Yes	No	Yes	No	No	Yes
62	Karatzas <i>et al.</i> <sup>99</sup>	[C] Possible malformations detection in <i>Daphnia</i> images and classification of daphnoids-[General methodology that can be used with different data and produce different models]	4323 images	No	Object detection: R-CNN coupled with a SVM model and SSD discretization method Classification: residual CNN and the ImageNet CNN	External validation	Yes	No	No	No	No	No
63	Robinson <i>et al.</i> <sup>72</sup>	[C] Lethality in embryonic zebrafish ( <i>Danio Rerio</i> ) expressed as classes of lethal or non-lethal NMs based on lowest observed effect levels (LOELs)	44	Yes	Random Forest, logistic regression	Internal and external validation	No	Yes	Yes	No	No	Yes

65	Gajewicz <sup>55</sup>	[R] Cytotoxicity to <i>E. Coli</i> (logarithmic values of molar 1/EC <sub>50</sub> ), Cytotoxicity to human keratinocyte cell line (HaCaT) expressed as the molar 1/log LC50	17, 18	Yes	kNN	External validation	Yes	Yes	No	Yes	No	Yes
66	Okse <sup>105</sup>	[R] Cell viability	10	No	PLS	Internal validation	No	No	Yes	No	No	No
67	Wang <i>et al.</i> <sup>107</sup>	[R] i. cellular uptake in human lung and ii. kidney cells (A549 and HEK293 cells), iii. ability to induce oxidative stress (indicated by the HO-1 level in the A549 cells), and iv. hydrophobicity (indicated by logP values).	41	Yes	kNN	Internal and external validation	No	Yes	Yes	Yes	No	Yes
68	Vijayaraghavan <i>et al.</i> <sup>81</sup>	[R] Tensile strength of HFG	31	No	ANN	External validation	No	No	No	No	No	Yes
69	Lazarovits <i>et al.</i> <sup>116</sup>	[R] i. Diameter, ii. Half-life, iii. Spleen accumulation (%ID) and iv. Liver accumulation (%ID)	120	Yes	Neural networks (feedforward and backpropagation cycle)	Internal validation	Yes	No	Yes	Yes	No	Yes
70	Mikolajczyk <i>et al.</i> <sup>142</sup>	[R] Zeta potential in cell culture media (serum-free)	15	Yes	GA-MLR, consensus modelling	Internal and external validation	Yes	Yes	Yes	Yes	No	Yes

71	Shao <i>et al.</i> <sup>69</sup>	[R] i. Protein binding in bovine serum albumin (BSA), ii. carbonic anhydrase (CA), iii. chymotrypsin (CT), and iv. hemoglobin (HB), v. nitrogen oxide (NO) generation, vi. cell viability by WST-1 assay and vii. multiassay score (arithmetic sum)	29	Yes	Genetic function approximation (variable selection)-MLR	Internal validation	No	No	Yes	Yes	No	No
72	Hassanzadeh <i>et al.</i> <sup>129</sup>	[R] Logarithm of adsorption coefficients (logk) of small organic compounds on MWCNTs, Logarithm of surface area normalized adsorption coefficients (logKSA), of organic compounds on MWCNTs	40, 69	Yes	whole space GA (genetic algorithm)-RBFN (radial basis function neural network)	Internal and external validation	Yes	Yes	Yes	No	No	Yes
73	Gonzalez-Durruthy <i>et al.</i> <sup>141</sup>	[R] Mitochondrial oxygen consumption (E3) (organism rat-liver)	16335	No	MLR, PLR, random forests (best performing), neural networks	External validation	No	No	No	Yes	No	Yes
74	Subramanian <i>et al.</i> <sup>133</sup>	[C] Cytotoxicity, encoded as "1" (true) if measured cell viability was <50% with respect to the control, and "0" (false) otherwise	483	Yes	Logistic regression, SVM, random forest, neural networks, consensus	Internal and external validation	No	Yes	Yes	Yes	No	Yes

75	Liu <i>et al.</i> <sup>66</sup>	[C] Toxicity class ("bioactive"/ "inactive") summarized after considering 64 bioactivity measures	44	Yes	LDA, Naïve Bayes Classifier (NBC), Logistic regression, and kNN	Internal validation	No	Yes	Yes	Yes	No	Yes
76	Kar <i>et al.</i> <sup>53</sup>	[R] Cytotoxicity to <i>E. Coli</i> (logarithmic values of molar 1/EC <sub>50</sub> )	17	Yes	MLR and PLS	Internal validation	Yes	Yes	Yes	Yes	No	Yes
77	Liu <i>et al.</i> <sup>40</sup>	[R] Cell association in A549 cell line (log <sub>2</sub> -transformed)	84	Yes	SVR regression, Linear regression	Internal validation	No	Yes	Yes	Yes	No	Yes
78	Liu <i>et al.</i> <sup>225</sup>	[R] Additive and weighted embryonic zebrafish metric (EZ Metric) (derived after consideration of different toxic endpoints in embryonic zebrafish), hpf (hours post-fertilization) scores.	656 (82 NMs)	Yes	kNN, KStar, MSP (a decision tree predictor), and Bagging (a bootstrapping meta-learning algorithm)	Internal validation	No	No	Yes	No	No	No
79	Borders <i>et al.</i> <sup>226</sup>	[R] Young's modulus and Poisson's ratio	78	Yes	PCA, Star plots, PLS regression	External validation	No	No	Yes	Yes	No	Yes

81	Furxhi <i>et al.</i> <sup>95</sup>	[C] Toxicity (toxic or nontoxic) via % cell viability	574, 144	No	Simple Decision Table majority classifier, BNs, Random Forests, kNN, Locally Weighted Learning, Neural Network Multilayer Perceptron, Sequential Minimal Optimization, Standard Linear Regression, Meta Vote Algorithm	Internal and external validation	Yes	No	Yes	No	No	Yes
83	Arts <i>et al.</i> <sup>125</sup>	[C]NMs classification in (1) soluble NMs, (2) biopersistent high aspect ratio NMs, (3) passive NMs, and (4) active NMs.	NA	No	None [Decision-making framework]	None reported	No	No	No	Yes	No	No
84	Ghorbanzadeh <i>et al.</i> <sup>46</sup>	[R] Cellular uptake in pancreatic cancer cells (PaCa2) (log-transformed)	109	Yes	SOM, multi-layered perceptron neural network (MLP-NN), MLR	Internal and external validation	Yes	Yes	Yes	Yes	No	No

85	Liu <i>et al.</i> <sup>156</sup>	[C] Summarized multi-perspective toxicity profile, from seven different assays for two different cell lines (human bronchial epithelial (BEAS-2B) cells and murine myeloid (RAW 264.7) cells). Toxicity class ("toxic"/ "non-toxic").	24	Yes	Classification using naive Bayesian classifier, linear regression, LDA, logistic regression, quadratic logistic regression and SVM	Internal validation	No	Yes	Yes	Yes	No	Yes
86	Guo <i>et al.</i> <sup>227</sup>	[R] Breakdown field strength of polyimide nanocomposite polyimide films	32	No	Sequential Minimal Optimization-Support Vector Regression (SMO-SVR), Stochastic Gradient Boosting, linear regression, back propagation neural network, generalized regression neural network, SVR	Internal validation	Yes	No	Yes	No	No	Yes
87	Sangani <i>et al.</i> <sup>228</sup>	[R] TiO <sub>2</sub> mobility in intact soil media (homogeneity expressed as mass recovery)	270	Yes	MLR, classification and regression tree (CART), random	Internal and external validation	No	No	Yes	Yes	No	No

					forest and ANN							
88	Timoshenko <i>et al.</i> <sup>229</sup>	[R] Average coordination number	18	No	ANN	External validation through different simulations (RMC-EXAFS simulations)	No	No	No	Yes	No	No
90	Wyrzykowski <i>et al.</i> <sup>98</sup>	[R] Zeta potential in ionized solution (KCl)	15	Yes	GA-MLR	Internal and external validation	Yes	Yes	Yes	Yes	No	Yes
91	Sizochenko <i>et al.</i> <sup>80</sup>	[C] Stability of NPs based on the zeta potential. Continuous zeta potential values were converted into labels: NPs with $\xi < -25$ mV and $\xi > 25$ mV were labelled as “stable”, remaining endpoints were labelled as “unstable”	208	Yes	Neural Network	Internal Validation	No	Yes	Yes	Yes	No	Yes
92	Marvin <i>et al.</i> <sup>230</sup>	[C] Hazard potential on various biological effects (none/low/medium/high) considering various effect/endpoint values	468	Yes	BNs	External validation	No	No	Yes	Yes	No	Yes
93	Gasper <i>et al.</i> <sup>231</sup>	[R] Adsorption energy (Eads) of a CO molecule on nanoPt clusters	195	Yes	Gradient-boosting regression algorithm	Internal validation	No	No	Yes	Yes	No	No

94	Jung <i>et al.</i> <sup>74</sup>	[R] Toxicity to <i>Daphnia Magna</i> and zebrafish ( <i>Danio rerio</i> ) expressed as -logLC <sub>50</sub> concentration	134, 98	Yes	Euclidean Distance grouping and MLR	It is stated that external validation is performed. However, the test data are not provided, and statistics seem to be derived by the training data.	No	No	No	Yes	No	No
95	De at al. <sup>62</sup>	[R] Cytotoxicity to <i>E. Coli</i> (logarithmic values of molar 1/EC <sub>50</sub> ), Cytotoxicity to human keratinocyte cell line (HaCaT) expressed as the molar 1/log LC <sub>50</sub> , Percentage decrease in enzymatic activity (enzyme inhibition to zebrafish in % (%EI_Zebrafish))	19,18,24	Yes	GA-MLR, PLS	Internal and external validation	Yes	Yes	Yes	Yes	No	Yes
96	Speck-Planche <i>et al.</i> <sup>232</sup>	[C] Antibacterial activity expressed as IC <sub>50</sub> , MIC, MBC, and Microb-Eff	300	Yes	Linear discriminant analysis (LDA), perturbation theory	External validation	Yes	No	No	Yes	No	No
98	Ramezanizadeh and Nazari <sup>128</sup>	[R] Thermal conductivity of Ag/water nanofluids	Not provided	No	MLR and GMDH ANNs	Training and validation using the same set	No	No	No	No	No	No



99	Findlay <i>et al.</i> <sup>84</sup>	[C] Protein corona formation (proteins enriched on Ag ENMs "PC" and proteins enrichment in solution "non-PC")	3012	Yes	Random Forest classification , two more algorithms were used as modelling approaches (LR and SVM) comparison with the existing RFC model	Internal validation	Yes	Yes	Yes	No	No	Yes
100	Kleandrov <i>et al.</i> <sup>233</sup>	[C] Ecotoxicity class of NPs by converting different toxicity measures (EC <sub>50</sub> , IC <sub>50</sub> , LC <sub>50</sub> , or TC <sub>50</sub> ) from assays on bacteria, algae, crustaceans, and fish to binary values according to cut-off values	5520	Yes	LDA	External validation	Yes	No	No	Yes	No	Yes
102	Toropov <i>et al.</i> <sup>132</sup>	[R] Mutagenic potential of MWCNTs expressed as the negative value of decimal logarithm of the TA100 (mean mutant counts after incubation of Salmonella strains)	24	Yes	Monte Carlo optimization	External validation (3 subsets)	Yes	Yes	Yes	No	No	No
103	Fernandez <i>et al.</i> <sup>83</sup>	[R] Graphene electronic properties: the energy of the band gap (EG), the ionization potential (EI), the energy of the Fermi level (EF), and	622	Yes	PLS Regression	Internal and external validation	Yes	No	Yes	No	No	No

		the electron affinity (EA).										
104	Basant and Gupta <sup>61</sup>	[R] Cytotoxicity in <i>E. Coli</i> and HaCaT cells	19,18,17 (single target), 16 (multi target)	Yes	Random forest regression	Internal and external validation	Yes	Yes	Yes	Yes	No	No
105	Roy <i>et al.</i> <sup>140</sup>	[R] Cytotoxicity to Chinese hamster ovary cells (logarithmic values of molar 1/EC <sub>50</sub> )	34	Yes	Linear regression based on MAE criteria, PLS	Internal and external validation	Yes	Yes	Yes	Yes	No	Yes
106	Alsharif <i>et al.</i> <sup>115</sup>	[R] Adsorption energies of human blood serum proteins	472 (59 proteins x 8 NMs)	Yes	ANN	External validation (3 subsets)	No	No	No	No	No	Yes
107	Trinh <i>et al.</i> <sup>130</sup>	[R] logEC <sub>50</sub> of TiO <sub>2</sub> based nano-mixtures, Immobilization (percentage of non-mobile and dead <i>D. magna</i> compared to the control sample of <i>D. magna</i> exposed to TiO <sub>2</sub> based nano-mixtures)	76, 313	No	MLR, random forest, SVM, and ANN	Internal and external validation	Yes	Yes	Yes	Yes	No	No
108	Brinkman <i>et al.</i> <sup>234</sup>	[R] Log k Adsorption Affinity of Metabolites to Metal and Carbon NMs	19	Yes	Biological surface adsorption index (BSAI) theory, MLR	Internal validation	No	Yes	Yes	Yes	No	No

109	Zhang <i>et al.</i> <sup>127</sup>	[R] Interaction energy (Eint)	17	Yes	Orthogonal PLS (OPLS) regression	Training and validation using the same set	No	No	Yes	No	No	Yes
110	El Yamani <i>et al.</i> <sup>126</sup>	[C] Toxicity by combining the results of three assays (cytotoxicity assessed via the AlamarBlue (AB) and colony forming efficiency (CFE) assays, and genotoxicity by the enzyme-linked version of the comet assay). Three classes are derived namely: (1) non-toxic (2) slightly toxic and (3) toxic	17	Yes	2D Hierarchical Clustering Analysis, PLS	None reported	No	No	No	Yes	No	Yes
111	Fanourgakis <i>et al.</i> <sup>235</sup>	[R] The adsorption capacity for CO <sub>2</sub> , H <sub>2</sub> , and H <sub>2</sub> S gases at specific thermodynamic conditions	2932	Yes	Random forest algorithm (RF)	Internal and external validation	No	No	Yes	Yes	No	No
112	Luan <i>et al.</i> <sup>96</sup>	[C] Cytotoxicity expressed the concentration of NPs leading to 50% reduction in cell viability (TC <sub>50</sub> ) (NPs with TC <sub>50</sub> >503.572 µM were considered non-cytotoxic, otherwise they were cytotoxic)	41	Yes	Linear discriminant analysis (LDA)	Internal and external (for Ni and SiO <sub>2</sub> NPs) validation	Yes	No	Yes	No	No	Yes

113	Jha and Yoon <sup>111</sup>	[C] Visual discrimination of toxic versus nontoxic NMs through PCA	73	Yes	PCA and SVM	External validation	Yes	No	No	No	No	Yes
114	Furxhi <i>et al.</i> <sup>135</sup>	[C] Nine different cellular effects in binary mode (Triggered, No effect)	245	Yes	BNs	Internal and external validation	No	No	Yes	Yes	No	Yes
115	Roy and Roy <sup>236</sup>	[R] Cytotoxicity (% cell viability) to human bronchial epithelial (BEAS-2B) and murine myeloid (RAW264.7) cell lines, expressed as effective concentration (log 1/EC <sub>50</sub> )	4 sub-datasets: 9, 10, 10, 12	Yes	PLS, simple linear regression and stepwise-MLR	Internal and external validation	Yes	Yes	Yes	Yes	No	Yes
116	To <i>et al.</i> <sup>139</sup>	[R] Toxicity in concentration-response in embryonic zebrafish (18 endpoints summed up using weighted Aggregate Entropy (wAggE) to characterize the severity of toxicity)	90 (15 NMs)	Yes	Random Forests decision tree (4 splits)	External validation	Yes	Yes	No	Yes	No	Yes
117	Fourches <i>et al.</i> <sup>68</sup>	[C] Protein binding to carbonic anhydrase (CA), cell viability by WST-1 assay, transformed to binary class. A separation bioactivity threshold at 2.0 was used to split the dataset in two classes of CNTs with balanced distribution CA binders (1) and CA	83	Yes	kNN, Random Forest, SVM	Internal and external validation	Yes	Yes	Yes	Yes	No	No

		non-binders (0). For the toxicity, a cut-off range was used [37–43%], splitting the dataset into toxic (1) and non-toxic (0) CNTs.										
118	Toropov <i>et al.</i> <sup>217</sup>	[R] Zeta potential in water	87	Yes	Simple linear regression - Monte Carlo optimization	External validation (3 subsets)	Yes	Yes	Yes	Yes	No	No
119	Toporov <i>et al.</i> <sup>237</sup>	[R] Cellular uptake in pancreatic cancer cells (PaCa2) (log-transformed)	109	No	Least squares method - Monte Carlo optimization	External validation (five random splits in 3 subsets)	Yes	Yes	Yes	No	No	No
120	Toporova <i>et al.</i> <sup>238</sup>	[R] Cytotoxicity - inhibition ratio (IR%) of SiO <sub>2</sub> NPs on human lung fibroblasts (HFL-Is)	16	No	Simple linear regression - Monte Carlo optimization	External validation (3 subsets)	No	No	Yes	No	No	No
121	Toporova <i>et al.</i> <sup>138</sup>	[R] Damage to cellular membranes (MD [units/L]) of immortalized human lung epithelial cells	42	No	Simple linear regression - Monte Carlo optimization	External validation (3 subsets)	Yes	Yes	Yes	No	No	No
122	Manganelli <i>et al.</i> <sup>131</sup>	[R] Viability (%) of cultured human embryonic kidney cells (HEK293), measured by the MTT assay	40	No	Simple linear regression - Monte Carlo optimization	External validation (3 subsets)	Yes	Yes	Yes	No	No	Yes
123	Toropov <i>et al.</i> <sup>239</sup>	[R] Mutagenicity expressed as negative decimal log transformed TA100	44	No	Simple linear regression - Monte Carlo optimization	External validation (3 subsets)	Yes	Yes	Yes	Yes	No	No

124	Yan <i>et al.</i> <sup>97</sup>	[R] i. nanohydrophobicity (experimental logP value that was defined by the ratio of the concentration of NPs in water to their concentration in octanol), ii. zeta potential in Millipore water with pH 7, iii. cellular uptake (adenocarcinomic human alveolar-based epithelial (A549) cells), and iv. protein adsorption (incubation in phosphate-buffered saline (PBS) medium)	147	No	Convolutional neural network (LeNet and GoogLeNet)	Internal and external validation	Yes	No	Yes	Yes	No	Yes
125	Wang <i>et al.</i> <sup>82</sup>	[R] Log-transformed surface area normalized adsorption coefficient (log K <sub>sa</sub> )	39 (aromatic compounds)	Yes	MLR, ANN, SVM	Internal and external validation	No	No	Yes	Yes	No	Yes
126	Baharifar and Amani <sup>240</sup>	[R] Cytotoxicity to Mrc-5 cells of chitosan/streptokinase NPs (Cs/SK NPs) (cell viability %) via MTT assay	30	No	ANN	External validation (3 subsets)	No	No	Yes	Yes	No	No
133	Baboukani <i>et al.</i> <sup>241</sup>	[R] "Friction": maximum energy barrier at the onset of sliding of a single layer 2D material against a similar layer in an inert environment	15	Yes	Bayesian Model	External validation (comparison to other studies and to simulations results)	No	No	Yes	Yes	No	Yes

134	Bilal <i>et al.</i> <sup>242</sup>	[R] % cell viability and IC <sub>50</sub> (mg L <sup>-1</sup> ) (in eukaryotic cell lines)	3028	Yes	BNs	Internal and external validation	No	No	Yes	Yes	No	Yes
136	Oh <i>et al.</i> <sup>243</sup>	[R] i. Cell viability: % of cell viability or other toxicity metric. ii. IC <sub>50</sub> Value: the exposure concentration at which there is 50% (or more) inhibition of cell growth or other toxicity metric.	1741	Yes	Random forest regression	Internal validation	No	No	Yes	Yes	No	Yes
137	Gernand and Casman <sup>244</sup>	[R] Pulmonary toxicity endpoints: i. polymorphonuclear neutrophils (PMNs) count, ii. macrophages (MAC) count, iii. lactate dehydrogenase (LDH) concentration, and iv. total protein (TP) concentration. These indicators were all measured in bronchoalveolar lavage (BAL) fluid extracted from the lungs of the mice or rats and were reported as a counts per subject or fold of control measurements.	136	Yes	Regression trees, random forest regression, stepwise linear regression	Internal validation	No	No	Yes	Yes	No	Yes
138	Choi <i>et al.</i> <sup>245</sup>	[C] Toxicity via Cell viability (%), NMs classified as "toxic" or "non-toxic"	574	No	Generalized linear model (Logistic GLM), SVM, Random	Internal and external validation	Yes	Yes	Yes	No	No	Yes

					Forest, Neural Network							
139	Bahl <i>et al.</i> <sup>114</sup>	[C] <i>In vivo</i> and <i>in vitro</i> categorization of the NMs, classified as "active" or "passive"	11	Yes	Random Forest, PCA-kNN approach	Internal validation	No	No	Yes	Yes	No	Yes
142	Romeo <i>et al.</i> <sup>109</sup>	[R] Viability of lung cells, Benchmark Dose Lower Confidence Interval (BMDL), [C] BMDL classification	1117 (see SI file)	Yes	Bayesian Ridge Regressor, Random Forest Regressor, XGB Regressor, Gradient Boosting Regressor, Multi-layer Perceptron Regressor, and Regression Voting Ensemble	Internal validation (nested LOO)	Yes	Yes	Yes	Yes	No	Yes
143	Sotiropoulos <sup>198</sup>	[C] Cytotoxicity class. NPs with a cell viability percentage over 50% were considered as toxic.	345	Yes	Random Forest classification	Internal and external validation	Yes	Yes	Yes	No	Yes	Yes
144	Gousiadou <sup>198</sup>	[R] Zeta potential in water	69	Yes	XGBoost Trees	Internal and external validation	No	Yes	Yes	No	Yes	Yes
146	Sotiropoulos <sup>198</sup>	[R] Toxicity expressed as median (%) DNA strand breaks	48	No	Extra-Trees Regressor, AdaBoost Regressor	Internal and external validation	Yes	Yes	Yes	No	Yes	No



148	Sotiropoulos <sup>198</sup>	[R] Cellular uptake capacity in A549 cells, logP and zeta potential in water	75 to 145 (not clear)	Yes	Extra - Trees Regressor, XGBoost Trees	Internal and external validation	No	Yes	Yes	No	Yes	No
153	Shin <i>et al.</i> <sup>201</sup>	[C] cytotoxicity class, [R] pBMD5, zeta potential	32 (BMD5), 42 (zeta pot.), 61 (cytotoxicity)	Yes	MLR, logistic regression	External validation	Yes	Yes	Yes	Yes	No	Yes
154	Ambure <i>et al.</i> <sup>199</sup>	[R] pEC <sub>50</sub> , membrane damage, smooth muscle apoptosis, solubility, bacterial reverse mutation test (NanoProfiler tool)	The tool hosts "ready-to-use" pre-trained models on different NM datasets	Yes	Linear regression	External validation	Yes	Yes	Yes	No	No	Yes

## Appendix II – Nanoinformatics models and methodologies available via GUIs

In this section we are briefly presenting the models and *in silico* methodologies found in Literature that are disseminated as web applications or stand-alone platforms. The key information of the models is presented in tabular format at the beginning of each paragraph, followed by a short presentation of their use. Emphasis is given in the availability of training material, considering that the absence of instructions renders the tools difficult to use. The models are presented in alphabetical order.

### A safe-by-design tool for functionalised nanomaterials: a web-service available through the Enalos Nanoinformatics Cloud Platform

NMs category:	Functionalised MWCNTs
Endpoint:	Protein binding of carbonic anhydrase & cell viability by WST-1 assay
Reliability:	Confusion matrix, accuracy, sensitivity, specificity
Domain:	Indication along with the prediction
Availability:	<a href="http://enaloscloud.novamechanics.com/EnalosWebApps/CNT/">http://enaloscloud.novamechanics.com/EnalosWebApps/CNT/</a>
Training material:	Written manual - <a href="http://enaloscloud.novamechanics.com/EnalosWebApps/CNT/instructions.zul">http://enaloscloud.novamechanics.com/EnalosWebApps/CNT/instructions.zul</a> YouTube video - <a href="https://youtu.be/BLuG-LwuQ1E">https://youtu.be/BLuG-LwuQ1E</a>

MWCNTs are currently used in numerous industrial applications and products, therefore fast and accurate evaluation of their toxicological and biological effect (protein binding of carbonic anhydrase) is of greatest importance. In this course, two predictive QNAR/kNN models for the assessment of decorated MWCNTs biological and toxicological profile are developed, validated, and released through a web-service<sup>70</sup>. The web-service provides to the user the possibility to import various data sets with organic MWCNT decorators of interest during a safe-by-design process. Therefore, their effects on the biological and toxicity behaviour of the resulting decorated MWCNTs can be assessed. To initiate the process, compound structures must be provided, either as SMILES notations list, via an SDF file, or simply as a structure drawing (Figure 70A). The predictions of the MWCNTs toxicity and biological activity, and their reliability are produced and presented to the user within seconds (Figure 70B). The neighbours' space of each decorating molecule can be also studied, given that their neighbours are also presented as a result to the users (Figure 70C). This web-service is an easily utilisable, user-friendly application aiming to facilitate decision making, by reducing expensive and time-consuming experimental procedures.

### A Enalos Nanoinformatics Cloud Platform: A Safe-by-Design Tool for Functionalised Nanomaterials

### B Safe-by-Design: Functionalised Nanomaterials

Knime report powered by Birt

"Toxicity"	"Domain (Toxicity)"	"Activity"	"Domain (Activity)"
toxic	reliable	non-binder	reliable
non-toxic	reliable	binder	reliable
toxic	reliable	non-binder	reliable
toxic	unreliable	binder	reliable

Date: Apr 2, 2019 5:43 PM  
www.knime.com

Author: NovaMechanics Ltd

1 of 1

### C

	A	B	C	D	E	F	G	H	I
	row ID	Activity	Neighbor 0	Distance of Neighbor 0	Neighbor 1	Distance of Neighbor 1	Neighbor 2	Distance of Neighbor 2	Domain (Activity)
2	Row0	non-binder	c1cccc(c1)C(Oc1cc	0.182235425	c1cccc(c1)C(Oc1cc	0.347628682	c1cccc(c1)C(Oc	0.360303746	reliable
3	Row1	binder	c1cccc(c1)C(Oc1cc	0.125957089	c1cccc(c1)C(Oc1cc	0.130089531	c1cccc(c1)C(Oc	0.144106101	reliable
4	Row2	non-binder	c1cccc(c1)C(Oc1cc	0.182235425	c1cccc(c1)C(Oc1cc	0.347628682	c1cccc(c1)C(Oc	0.360303746	reliable
5	Row3	binder	c1cccc(c1)C(Oc1cc	0.112993142	c1cccc(c1)C(Oc1cc	0.123445261	c1cccc(c1)C(Oc	0.154328932	reliable

Figure 70: [A] Screenshot of the Enalos cloud platform safe-by-design tool for functionalised MWCNTs interface. [B] Results table including toxicity and biological activity predictions and their reliability, for each input compound. [C] Neighbours space for each input compound.

### Apellis

NMs category:	General methodology for models' generation (NMs type based on input data)
Endpoint:	General methodology for models' generation (endpoint based on input data)
Reliability:	External explained variance, MSE, MAE, confusion matrix, MCC, accuracy, sensitivity, specificity
Domain:	If no neighbours are identified for query NM, it is located outside the DoA of the developed model
Availability:	<a href="https://apellis.jaqpot.org/">https://apellis.jaqpot.org/</a>
Training material:	Written manual – <a href="https://apellis.jaqpot.org/Apellis%20user%20guide.pdf">https://apellis.jaqpot.org/Apellis%20user%20guide.pdf</a> YouTube video – <a href="https://youtu.be/kjLXOAOiafs">https://youtu.be/kjLXOAOiafs</a>

Apellis is a web-application for read-across model development. Users can train a predictive model (for any properties-endpoint dataset), download it, and use it to acquire predictions of NM-related endpoints either numerical or categorical. Apellis incorporates the concepts of grouping, read-across, and mathematical optimization through genetic algorithms<sup>34,123</sup>. Users can upload their data in tabular format using a CSV file (templates are available) and tune the necessary parameters from the respective fields through menus, radio buttons, etc. During the training procedure, the application selects the most important NM properties that affect the NMs endpoint (variable selection). In the process of grouping NMs for performing read-across predictions, their multi-perspective characterization can be considered, by defining more than one similarity criteria (thresholds). The workflow converges to the

The screenshot displays the Apolite web application interface, which is used for evaluating machine learning models. The interface is divided into several sections:

- Top Navigation Bar:** Includes links for HOME, NUMERICAL SINGLE OUTPUT, NUMERICAL MULTIPLE OUTPUT, CLASS SINGLE OUTPUT, and CLASS MULTIPLE OUTPUT.
- Left Sidebar:** Contains options for EXPERIMENTATION and PERFORMANCE, a Choose training file dropdown (set to Use demo dataset), a Train button, and checkboxes for Scaling of raw data, Partitioning method (Random Split), Random, and Training ratio (0.6).
- Main Content Area:**
  - Experimental vs. predicted endpoint values for the two sets:** A scatter plot showing Predicted values (Y-axis, -8 to 3) versus Experimental values (X-axis, -9 to 3). Data points are colored red for 'test' and blue for 'training'.
  - Subset:** A table showing the distribution of data points for the 'test' and 'training' subsets.
  - Model Summary:** A table providing key performance indicators (KPIs) for the model, including Accuracy, Precision, Recall, F1 Score, and AUC.
  - Model Performance:** A table showing the model's performance across different metrics (Accuracy, Precision, Recall, F1 Score, AUC) for both 'test' and 'training' sets.
  - Model Evaluation:** A table showing the model's performance across different metrics (Accuracy, Precision, Recall, F1 Score, AUC) for both 'test' and 'training' sets.
  - Model Training:** A table showing the model's performance across different metrics (Accuracy, Precision, Recall, F1 Score, AUC) for both 'test' and 'training' sets.
- Right Sidebar:** Contains a table of model performance metrics (Accuracy, Precision, Recall, F1 Score, AUC) for both 'test' and 'training' sets.

The interface is designed to provide a comprehensive overview of the machine learning model's performance, allowing users to compare the model's performance across different metrics and datasets.

BioDaph

NMs category:	TiO <sub>2</sub> NMs
Endpoint:	TiO <sub>2</sub> concentration internalised by <i>Daphnia magna</i>
Reliability:	Comparison to experimental studies
Domain:	The user is advised to upload scenarios for TiO <sub>2</sub> concentration within or near the range of concentration values that the model of Fan <i>et al.</i> (2016) <sup>246</sup> was built upon, e.g., 0.1-10 mg/L.
Availability:	<a href="https://biodaph.cloud.nanosolveit.eu/">https://biodaph.cloud.nanosolveit.eu/</a>
Training material:	Written tutorial – available through the app

183

water concentration over time) and the simulation parameters (the time increment, the simulation period, and the type of TiO<sub>2</sub> NM). After the simulation, the TiO<sub>2</sub> concentration internalised by *D. magna* -time profile is presented in graph format.

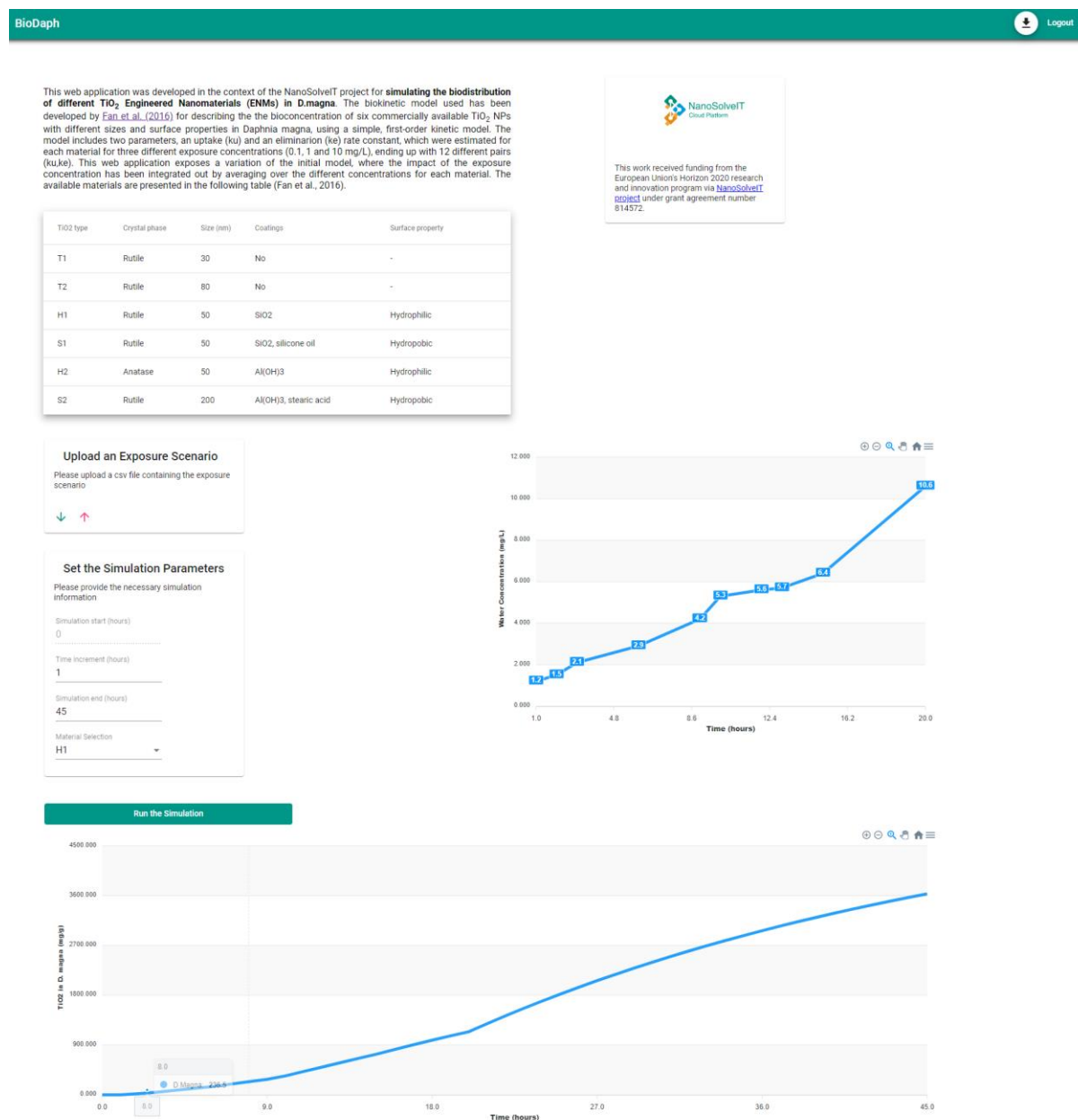



Figure 72: The NanoSolveIT MeOx BioDaph interface in Jupyter. Users provide their data by uploading a CSV file with the exposure scenario and type the simulation parameters. The simulated bioconcentration in *Daphnia Magna* time profile is presented at the bottom of the interface.

## Cytotoxicity (cell viability) prediction for metal oxide NPs

NMs category:	MeOx NPs
Endpoint:	Cytotoxicity to BEAS-2B and RAW 264.7 cell lines
Reliability:	Coefficient of determination ( $R^2$ ), external explained variance, Golbraikh & Tropsha's test
Domain:	Indication along with the prediction
Availability:	<a href="https://cellviability.cloud.nanosolveit.eu/">https://cellviability.cloud.nanosolveit.eu/</a>
Training material:	Written tutorial - <a href="https://cellviability.cloud.nanosolveit.eu/instructions.zul">https://cellviability.cloud.nanosolveit.eu/instructions.zul</a>

A web-application is provided through the NanoSolveIT Cloud Platform, which evaluates the cytotoxicity effects of MeOx NPs on human bronchial epithelial (BEAS-2B) and murine myeloid (RAW 264.7) cell lines for two types of cytotoxicity assays: the single parameter ATP and the LDH. A  $k$ NN model is provided that predicts the cytotoxicity of MeOx NPs expressed as the cell viability percentage<sup>155</sup>. The following set of descriptors used in the model are also the indicated properties required as input for the online tool: NP core size, NP hydrodynamic size, assay type, exposure dose, energy of the MeOx conduction band ( $E_c$ ), the coordination number of metal atoms in the NP (Coord. #Me atoms) and the force vector surface normal component of atoms ( $V_{\perp}$  #all atoms). Input NPs' data can be provided either by filling the provided form or by uploading a CSV file -following a given template (Figure 73). The results include the predicted cell viability % for each query NP, the identities, and distances of the closest neighbours on which the prediction was based, and an indication of whether the prediction can be considered reliable based on the model's DoA. The results can be downloaded in CSV format.

 **NanoSolveIT Cytotoxicity (Cell Viability) Prediction for Metal Oxide NPs**

[User Guide](#) A

User row ID	Core size (nm)	Hydro size (nm)	$E_c$ (eV)	Assay	Exp. dose (ug/mL)	Coord. #Me atoms	$V_{\perp}$ #all atoms
CuO	25	45	-5.17	ATP	3.2	3.579	-0.243
CeO2	64	120	-3.8	LDH	0.8	7.011	-1.108
HfO2	14	78	-2.96	ATP	200	6.215	-0.924
ZnO	89	92	-3.89	ATP	50	3.586	0.868
				ATP			

[Download files](#) B

ID	APD Prediction	kNN Prediction	Closest NN1	Distance from NN	Closest NN2	Distance from NN	Closest NN3	Distance from NN	Closest NN4	Distance from NN
CuO	reliable	1.6576248278753	ZnO	2.2228655489475	ZnO	2.2232594323735	ZnO	2.2275501962159	ZnO	2.2491214514422
CeO2	reliable	1.9480636756013	In2O3	2.6343319087030	In2O3	2.6343620830066	In2O3	2.6346034649939	In2O3	2.6357577383235
HfO2	reliable	1.9494421379711	La2O3	1.6744372048920	La2O3	2.2995194093385	NiO	2.5362383156147	SiO2	2.6053928616888
ZnO	unreliable	1.8548693934997	CoO	2.7147474904932	CoO	2.7431928540824	CoO	2.7783400451756	CoO	2.8131691040787

Figure 73: [A] Cell viability web service for the prediction of MeOx NP cytotoxicity. The user inputs the required parameters, and the calculation is performed automatically. [B] The results provided from the web service include the prediction, its reliability and the neighbours' space for each input NP.

## DeepDaph

NMs category:	TiO <sub>2</sub> , Ag, or AgS NMs
Endpoint:	Recognition of head, eye, heart, abdomen/claw, tail, tail-tip, and tail-base regions of the <i>Daphnia</i> and their overall length and tail length
Reliability:	Confusion matrix, accuracy, sensitivity, specificity, MCC
Domain:	Not provided
Availability:	<a href="https://deepdaph.cloud.nanosolveit.eu/">https://deepdaph.cloud.nanosolveit.eu/</a>
Training material:	Not available

The DeepDaph web application employs a deep learning workflow based on neural networks architecture to detect, isolate, analyse, and classify specific areas of interest of *Daphnids* depicted in microscopy images<sup>99</sup>. In this way, it is possible to spot malformations of the depicted organisms in a fast and systematic manner. Users should upload a PNG or JPEG image of *Daphnia* light microscopy and the required meta-data (generation, age in days, NM type of exposure, etc.). When submitted the deep learning workflow is executed and the results are presented shortly, including the recognition of seven regions of interest (head, eye, heart, abdomen/claw, tail, tail-tip, and tail-base), the overall length and tail length of the *Daphnia*, and the classification of the heart and the abdomen/claw areas in terms of lipid concentration (Figure 74). This application is hosted in the NanoSolveIT cloud platform.

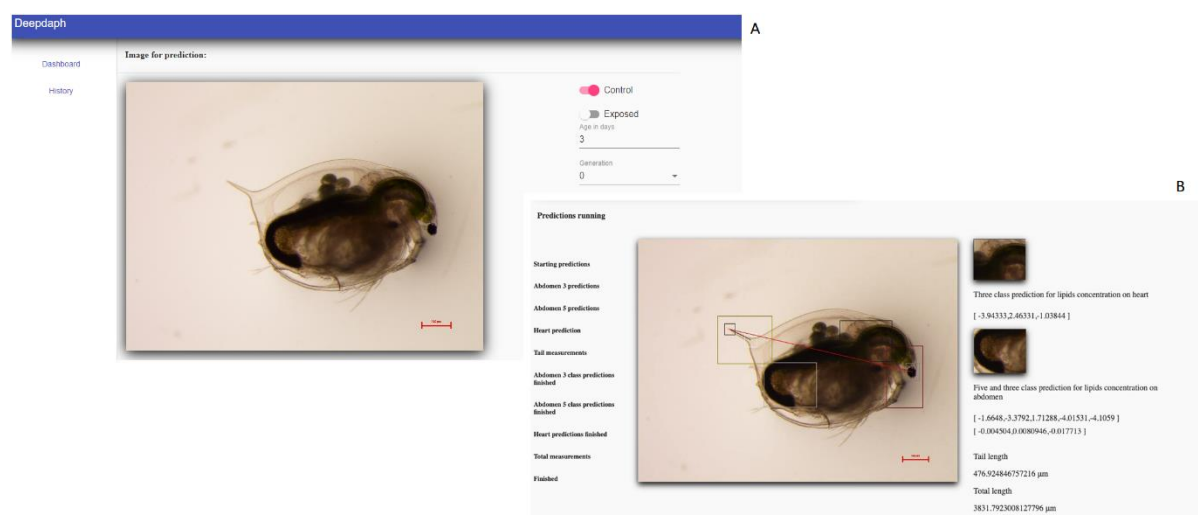



Figure 74: DeepDaph web service interface. [A] *Daphnia* image and meta-data input. [B] Results of the deep learning workflow.

## Ecotox models

NMs category:	Freshly dispersed and 2-year aged metal and MeOx NMs
Endpoint:	Toxicity to <i>Daphnia Magna</i>
Reliability:	Confusion matrix, accuracy, sensitivity, specificity, MCC
Domain:	Indication along with the prediction
Availability:	<a href="https://ecotox.cloud.nanosolveit.eu/">https://ecotox.cloud.nanosolveit.eu/</a>
Training material:	Written tutorial - <a href="https://ecotox.cloud.nanosolveit.eu/instructions.zul">https://ecotox.cloud.nanosolveit.eu/instructions.zul</a>

The toxicological effects on *Daphnia Magna* of freshly dispersed and 2-year aged NMs, can be assessed using the NanoSolveIT Cloud Platform ecotox models<sup>75</sup>. Two different kNN classification models -one for each ageing profile- are developed considering *Daphnia*'s EC<sub>40</sub> toxicity values. After choosing the


**NanoSolveIT Ecotox Model**

---

Age Profile

>

User Row ID	Core	Charge	Tested media	TEM [nm]	DLS [nm]	Concentration [ppb]
	Ag	positive	HH combo			
	Ag	positive	HH combo			
	Ag	positive	HH combo			
	Ag	positive	HH combo			
	Ag	positive	HH combo			
	Ag	positive	HH combo			
	Ag	positive	HH combo			
	Ag	positive	HH combo			
	Ag	positive	HH combo			
	Ag	positive	HH combo			
	Ag	positive	HH combo			
	Ag	positive	HH combo			
	Ag	positive	HH combo			
	Ag	positive	HH combo			
	Ag	positive	HH combo			
	Ag	positive	HH combo			
	Ag	positive	HH combo			
	Ag	positive	HH combo			
	Ag	positive	HH combo			
	Ag	positive	HH combo			
	Ag	positive	HH combo			

Execute computations
Reset

Upload csv file
CSV title

Execute computations
Reset

A

B

## Enalos QNAR iron oxide toxicity platform

NMs category:	Iron oxide NPs
Endpoint:	Toxicity class
Reliability:	Confusion matrix, accuracy, sensitivity, specificity, precision
Domain:	Indication along with the prediction
Availability:	<a href="http://www.enaloscloud.novamechanics.com/EnalosWebApps/QNAR_IronOxide_Toxicity/">http://www.enaloscloud.novamechanics.com/EnalosWebApps/QNAR_IronOxide_Toxicity/</a>
Training material:	Written tutorial - <a href="http://www.enaloscloud.novamechanics.com/EnalosWebApps/QNAR_IronOxide_Toxicity/instructions.zul">http://www.enaloscloud.novamechanics.com/EnalosWebApps/QNAR_IronOxide_Toxicity/instructions.zul</a> YouTube video: <a href="https://youtu.be/8Rxo_dhmD34">https://youtu.be/8Rxo_dhmD34</a>

187



toxicity is developed on KNIME platform using the J48 modeling methodology and is available online via Enalos Cloud Platform<sup>65</sup>. Through this web service (Figure 76), the user can insert the indicated properties (NP size, zeta potential, relaxivities R1 and R2) and choose among three alternative coating options (Poly-vinyl-alcohol-PVA, cross-linked dextran or other) for multiple NP entries. Input properties can be provided either by filling in the given form or by uploading a CSV file. Afterwards, a toxicity prediction is generated, classifying each input NP as “active” or “inactive”, accompanied by an indication of its reliability based on the DoA. There is an additional option to download output results in different formats (CSV or HTML file).

**A**

**Enalos QNAR Iron Oxide Toxicity Platform**

[User Guide](#)

MNP Numb	Size (nm)	ZP (mV)	R1 (mM-1S-1)	R2 (mM-1S-1)	Coating
1					Other
2					Other
3					Other
4					Other
5					Other
6					Other
7					Other
8					Other
9					Other
10					Other
11					Other
12					Other
13					Other
14					Other
15					Other
16					
17					
18					
19					
20					

[Download files](#)

[Execute computations](#)

**B**

Row ID	Prediction	Domain
"Row1"	"inactive"	"reliable"
"Row2"	"inactive"	"reliable"
"Row3"	"inactive"	"reliable"
"Row4"	"inactive"	"reliable"
"Row5"	"inactive"	"reliable"
"Row6"	"inactive"	"reliable"
"Row7"	"active"	"reliable"
"Row8"	"active"	"reliable"
"Row9"	"active"	"reliable"
"Row10"	"active"	"reliable"
"Row11"	"active"	"reliable"
"Row12"	"inactive"	"reliable"
"Row13"	"inactive"	"reliable"

Figure 76: [A] Screenshot of the iron oxide toxicity web interface. [B] Results table for each input NM including the toxicity predictions and their reliability.

### Facet cytotoxicity prediction

NMs category:	MeOx NPs
Endpoint:	Toxicity to BEAS-2B and RAW 264.7 cell lines and <i>E. Coli</i>
Reliability:	Cohen's kappa, accuracy, sensitivity, specificity, MCC
Domain:	Indication along with the prediction
Availability:	<a href="https://facetcytotoxicity.cloud.nanosolveit.eu/">https://facetcytotoxicity.cloud.nanosolveit.eu/</a>
Training material:	Written tutorial - <a href="https://facetcytotoxicity.cloud.nanosolveit.eu/instructions.zul">https://facetcytotoxicity.cloud.nanosolveit.eu/instructions.zul</a>

Risk assessment techniques associated with MeOx NPs are on the rise, due to the immediate need to determine NPs as safe to use health-wise or not. The NanoSolveIT facet cytotoxicity ready-to-use web

application uses a QSAR model to predict the effect of MeOx NP facets on their toxicity towards bronchial epithelial (BEAS-2B), murine myeloid (RAW 264.7) cell lines, and *E. coli* (Figure 77). Specifically, a kNN algorithm is implemented to classify NPs as “toxic” or “non-toxic” based on five significant parameters, namely core size, chemical potential, enthalpy of formation, electronegativity count (Metal Epsilon) and NP’s core size<sup>88</sup>. These properties are used as input data requirements for the online tool and users may provide them through the online form or by uploading a CSV file with a specific format. Computations are executed and besides NP’s cytotoxicity classification, output results contain the neighbouring NPs’ identities and Euclidean distances, along with a warning about the DoA of the prediction.

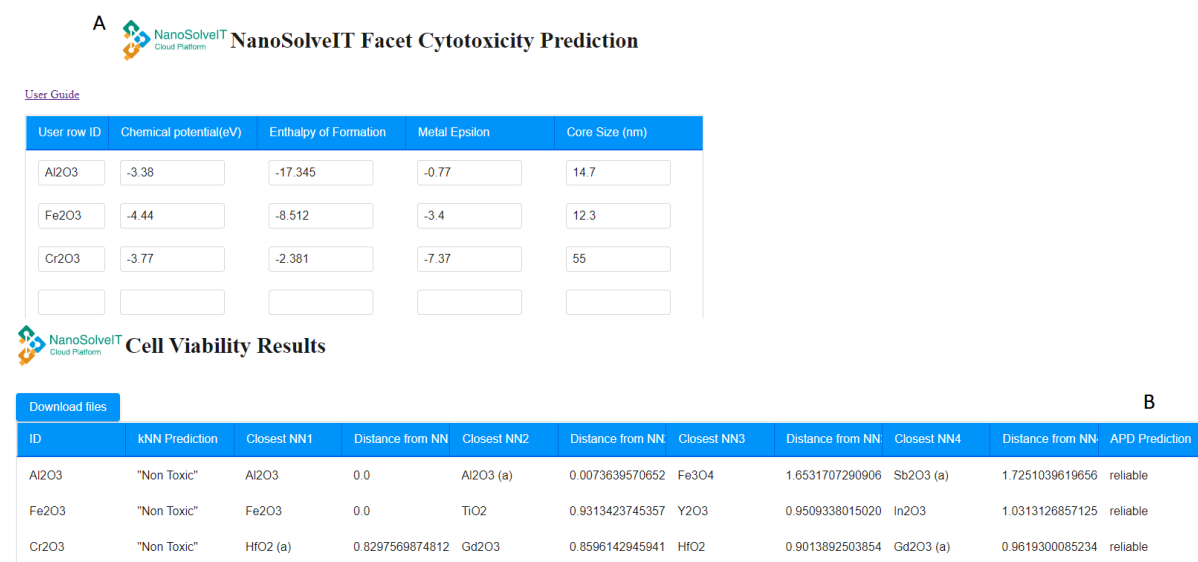


Figure 77: [A] NanoSolveIT facet cytotoxicity predictive model web service. [B] Results table including cytotoxicity class predictions, their reliability and the neighbours’ space of the input NMs.

## GUIDEnano

NMs category:	User-defined
Endpoint:	Human and environmental safety assessment (various endpoints)
Reliability:	Not provided
Domain:	Not provided
Availability:	<a href="https://tool.guidenano.eu/">https://tool.guidenano.eu/</a>
Training material:	YouTube video - <a href="https://youtu.be/sKMkr0-p3l0">https://youtu.be/sKMkr0-p3l0</a>

GUIDEnano is a web-based guidance tool, aiming to develop innovative methodologies to evaluate and manage human and environmental health risks of nano-enabled products, thus helping the industry to apply risk management strategies for a product of interest. The users are required to create an account for access to the tool, allowing them to describe their case study in terms of activities (>200 processes), materials (NM properties and other substances), compartments/fate, exposure, and hazard-risk assessment, by utilising the tools’ interconnected modules. GUIDEnano includes several predictive models, multi-level decision trees and databases, while it simulates a NM’s fate and provides estimates in human exposure routes and target environmental scenarios. The release values and occupational exposure limits can be either input by the user or derived from the tool from different study templates. Finally, the tool returns a risk assessment result after considering the whole product life cycle, and indicates the risk level as “high”, “medium”, or “low” (Figure 78). The users can also use the GUIDEnano

tool to simulate how risk predictions may change by introducing different types of risk mitigation measures.

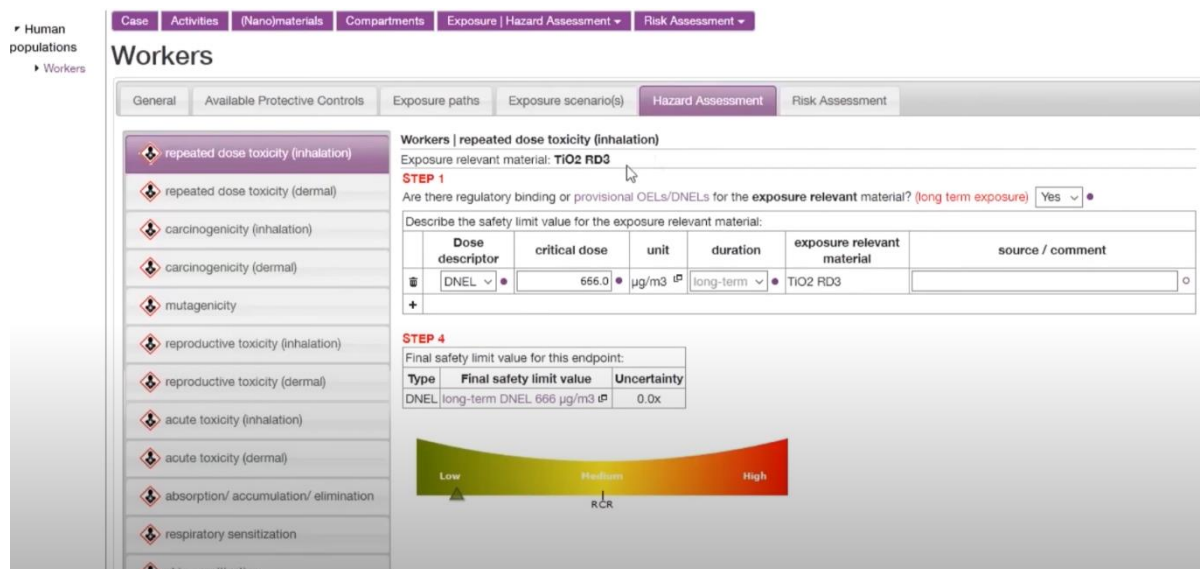


Figure 78: Screenshot of GUIDEnano interface.

## INSIdE nano

NMs category:	Metal and MeOx NMs, and MWCNTs
Endpoint:	Assessment of the connections between phenotypic entities based on their effects on the genes
Reliability:	Comparison with other independently calculated similarities based on several aspects, including the molecular composition of the drugs and chemicals, the symptoms of the diseases, the use of drugs in clinical practice, and the pathogenic effects of chemical exposures.
Domain:	Not provided
Availability:	<a href="http://inano.biobyte.de/">http://inano.biobyte.de/</a>
Training material:	Written tutorial - <a href="http://inano.biobyte.de/help.cgi">http://inano.biobyte.de/help.cgi</a>

The INSIdE NANO web tool depicts and highlights connections between phenotypic entities based on their effects on genes, for four categories of elements (NMs, drugs, chemicals and diseases)<sup>167</sup>. The tool provides two approaches for the exploitative analysis of the data set: the visualization of the phenotypic network and the visualization clustering of the phenotypic entities (Figure 79). For the NMs, the network construction is based on an ordered list of genes resulting from differential expression analysis. Pairwise similarity is then calculated between all entities and the respective values are used to build a weighted undirected network, where the nodes are the entities and the similarity between them represents the weight of the edge. Positive and negative correlations between nodes-entities are encoded in the colour of the edges. The cluster analysis panel gives the possibility to investigate how the four entities are grouped with each other and which genes are most important to each of them. In addition, query analysis (simple or conditional) is possible through the tool and allows the users to examine connections of a specific element and highlights its position in the ranking of the neighbours, it indicates if the connection is known in the literature and specifies the connection distributions.

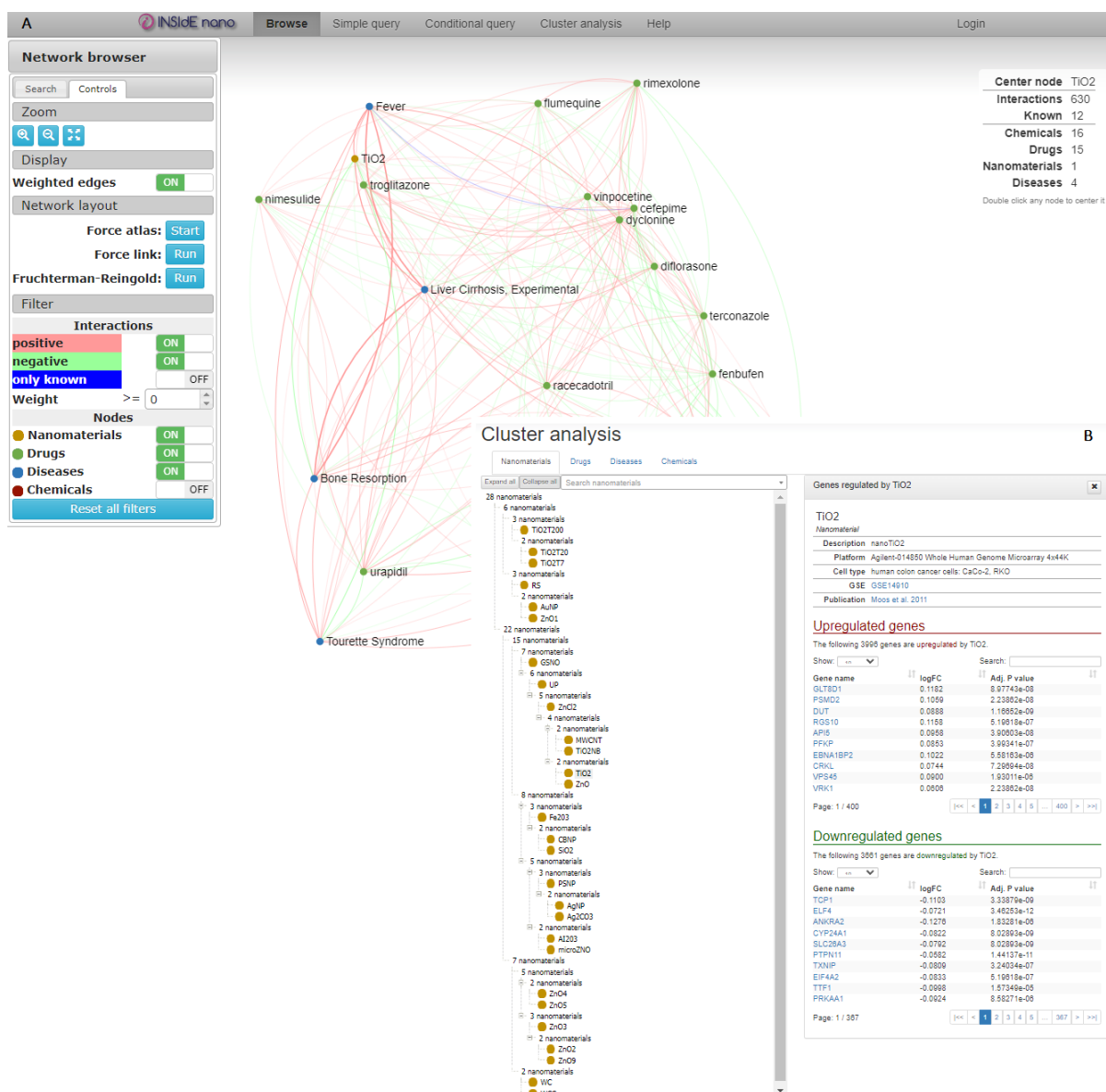


Figure 79: The INSIDE nano web tool. [A] The network browser tab allows the visualization of interactions between entities (NMs, drugs, chemicals, and diseases). [B] The cluster analysis tab allows the investigation of how NMs, drugs, disease and chemicals are grouped and of which genes are most important to each of them.

## Logistic regression model for the toxicity classification of SPIONs

NMs category:	SPIONs
Endpoint:	Cell viability of stem cells
Reliability:	Accuracy, precision, recall, F1-score, confusion matrix
Domain:	Assessed but not provided within the tool
Availability:	<a href="https://app.jaqpot.org/model/DcWnWFp9GESI16R4o2av">https://app.jaqpot.org/model/DcWnWFp9GESI16R4o2av</a> (upon request to become a member of the BIORIMA organization from hsarimv[AT]central[DOT]ntua[DOT]gr)
Training material:	Not available

A nanoQSAR model is developed for the cell viability of stem cells prediction of SPIONs and released as web service through the Jaqpot 5 modelling platform<sup>76</sup>. The developed logistic regression model

classifies input SPIONs as “toxic” and “non-toxic” (class 1 and 0 respectively), based on the values of two input features: the SPIONs type of core (maghemite or magnetite) and their overall size. Users after providing the values of these features to the web service, can acquire the predicted genotoxicity values (Figure 80).

**A**

**Jaqpot**

**MODEL**  
 Title: Logistic Regression model for the toxicity classification of SPIONs  
 Owner: Haralambos Sarimis  
 Description: Logistic Regression Model

**Logistic Regression model for the toxicity classification of SPIONs.**  
[Training dataset is available in this link](#)  
[Test dataset is available in this link](#)

**QMRf Report**

**1. QSAR Identifier**

**1.1.QSAR identifier (title):**  
 Logistic Regression model for predicting cell viability of superparamagnetic iron oxide nanoparticles (SPIONs) in stem-cell monitoring.  
 The model has been presented in the publication “QSAR modeling of the toxicity classification of superparamagnetic iron oxide nanopartic integrated study from data curation to model development”, M. Kotzabasaki, I. Sotiriopoulos, H. Sarimis, RSC Adv., 2020, 10 (9), 5385-53.

**1.2.Other related models:**  
 None

**1.3.Software coding the model:**  
 Jaqpot is a web platform that support development, validation and sharing of QSAR models [jaqpot.org](https://jaqpot.org)

**2. General Information**

**2.1.Date of QMRf:**  
 28 September 2019

**2.2.QMRf author(s) and contact details:**  
 Iason Sotiriopoulos, National Technical University of Athens, [iason.sotiri1@gmail.com](mailto:iason.sotiri1@gmail.com)  
 Marianna Kotzabasaki, National Technical University of Athens, [mariannako@chemeng.ntua.gr](mailto:mariannako@chemeng.ntua.gr)

**2.3.Date of QMRf update(s):**  
 28 September 2019

**2.4.QMRf update(s):**

**B**

**Dependent feature / Predicted feature**

**cell viability**  
 Description: Cell viability is the toxicological endpoint of this model and is defined as the quantification of the number of live cells and is usually expressed as a percentage of the control.  
 Units: 0 (Non-Toxic) - 1 (Toxic)

**Independent features**

**Overall size (nm)**  
 Description: Average size of the SPIONs  
 Units: 2-150 (nm)

**Magnetic core**  
 Description: Categorical feature separating maghemite ( $\gamma$ -Fe<sub>2</sub>O<sub>3</sub>) from magnetite (Fe<sub>3</sub>O<sub>4</sub>)  
 Units: 0: Fe<sub>3</sub>O<sub>4</sub>, 1:  $\gamma$ -Fe<sub>2</sub>O<sub>3</sub>

**C**

**Choose method**

Predict

**Upload dataset with the required independent features and values**

↓ ↑

**Input values for the independent features**

Sort

Magnetic core  
0  
0: Fe<sub>3</sub>O<sub>4</sub>, 1:  $\gamma$ -Fe<sub>2</sub>O<sub>3</sub>

Overall size (nm)  
47  
2-150 (nm)

**D**

**Task started**

Predicted dataset or value will be returned

PREDICTION Task is now running.  
 Prediction Task is now running.  
 Model retrieved successfully.  
 Searching dataset...  
 Dataset has been retrieved.  
 Starting Prediction...  
 Prediction completed successfully.  
 Dataset was built successfully.  
 Now saving to database...  
 Task Completed Successfully.

**View predicted value only**

Id	Overall size (nm)	Magnetic core	cell viability
0	47	0	0

Items per page: 30 1 - 1 of 1 < >

Figure 80: [A] Jaqpot logistic regression model for the toxicity classification of SPIONs main interface. [B] Information on the necessary input variables and the predicted endpoint. [C] Prediction page where users upload their data in CSV format or via the provided online form. [D] Model output.

## Lung exposure dose calculator

NMs category:	Black toner, TiO <sub>2</sub> and embedded TiO <sub>2</sub> NPs (TiO <sub>2</sub> -Ag, TiO <sub>2</sub> -Ag <sub>2</sub> O and TiO <sub>2</sub> -AgCl)
Endpoint:	NMs lung burden following acute exposure
Reliability:	Comparison to MPPD model results
Domain:	Provided only for the multi-box aerosol method
Availability:	<a href="https://lungexposure.cloud.nanosolveit.eu/">https://lungexposure.cloud.nanosolveit.eu/</a>
Training material:	Written tutorial - <a href="https://lungexposure.cloud.nanosolveit.eu/instructions.zul">https://lungexposure.cloud.nanosolveit.eu/instructions.zul</a>

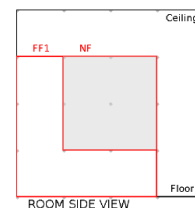
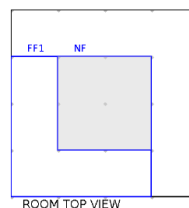
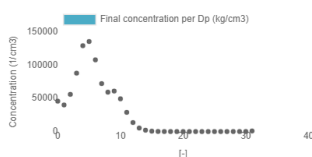
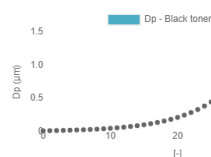
A biokinetics model for the estimation of acute regional lung-deposited dose of inhaled inorganic engineered NPs in humans is provided as a web application through the NanoSolveIT Cloud Platform<sup>161</sup>. The first step for the user is to provide data for the concentration of NPs over time using three alternative models for NP diameter distribution (theoretical, template, or from the multi-box aerosol model also hosted in NanoSolveIT Cloud). Next, two models can be used for the calculation of the NP deposition in the human respiratory system: the international commission on radiological protection model (ICRP) and an advanced ICRP model. Users must provide detailed parameter values for the exposed person in case that the advanced ICRP model is selected. The web-service, then, calculates the acute doses (in mg) for three regions of the respiratory system (alveolar, tracheobronchial, and head airways), which can finally be downloaded in CSV format (Figure 81).

[User Guide](#)

Select method: From aerosol ...

Aerosol configuration

Select area: NF - Near field



Select model: Advanced ICRP

Person exposed	Adult female
Dominant breathing type	Mouth
Activity level	Heavy exercise
Windspeed relative to the target for modulating inhalable fraction (up to 10 m/s) (m/s)	0
Functional residual capacity (-)	2.681
Dead volume upper airways (-)	0.04
Dead volume bronchi (-)	0.04
Dead volume Bronchioles (-)	0.044
Scaling factor of upper airways with adult male (-)	1.08
Scaling dimensions of bronchi and bronchioles with adult male (-)	1.04
Scaling dimensions of alveoli with adult male (-)	1.07
Tidal volume of exposed subject (L)	1.364
Exposure duration (sec)	2000
Volumetric flow rate of inspired air (mL/s)	1500
Density (Kg/cm³)	1000

Calculate dose

Calculated doses

	Alveolar	Head airways	Tracheobronchial
Acute dose (mg)	0.210241	0.098741	0.295784

Download doses

Figure 81: NanoSolveIT lung exposure dose calculator web interface. Users should define the particle size distribution and according to the selected model, the calculated acute doses in the alveolar, tracheobronchial and head airways are presented.

## Metal oxide nanoparticles cytotoxicity classification

NMs category:	MeOx NPs
Endpoint:	Cytotoxicity class
Reliability:	Accuracy, specificity, sensitivity, MCC, confusion matrix
Domain:	Assessed but not provided within the tool
Availability:	<a href="https://cytotoxicity.cloud.nanosolveit.eu/">https://cytotoxicity.cloud.nanosolveit.eu/</a>
Training material:	Written tutorial – available through the app

Another web service that provides cytotoxicity predictions for MeOx NPs is made available through Jaqpot Services, developed as part of the NanoSolveIT EU programme. A fully validated predictive model developed using the random forest classifier based on physicochemical and atomistic properties is developed. Users can easily access the service by creating a free Jaqpot account or via Google/GitHub accounts. Users should provide the required variable values either by completing the provided form or by uploading a CSV file following the provided template file. Results are presented shortly after the submission of the input data and include the predicted toxicity class (Figure 82).

### Metal Oxides Cytotoxicity Classification

Logout

Add row Submit

ID	Exposure dose (ug/mL)	Ec (eV)	Vj Me atoms surface	Total surface area (m <sup>2</sup> )	Vj Me atoms all	
CuO	200	-5.17	1.427531	14.18	0.309687	Remove
CoO	200	-4.42	1.217621	2.598	0.050290965	Remove
ZnO	200	-3.89	0.756355	9.356	0.153232	Remove
						Remove
						Remove
						Remove
						Remove
						Remove
						Remove
						Remove

#### Predict via CSV file


Upload dataset with the required independent features and values.

↓ ↑

#### Documentation

Download the QMRFP Report of the model and a tutorial on how to use this app.

↓



This work received funding from the European Union's Horizon 2020 research and innovation program via [NanoSolveIT project](#) under grant agreement number 814572.

ID	Toxicity
CuO	Toxic
CoO	Toxic
ZnO	Toxic

Download Predictions

Figure 82: NanoSolveIT MeOx cytotoxicity classification interface in Jaqpot. Users provide their data either manually or by uploading a CSV file. Predictions are presented at the bottom of the interface.



### MS<sup>3</sup>bD zeta potential predictive model

NMs category:	Metal and MeOx NMs
Endpoint:	Zeta potential in water (pH 7)
Reliability:	Coefficient of determination ( $R^2$ ), external explained variance, Golbraikh & Tropsha's test
Domain:	Indication along with the prediction
Availability:	<a href="http://www.enalosccloud.novamechanics.com/nanocommons/mszeta/">http://www.enalosccloud.novamechanics.com/nanocommons/mszeta/</a>
Training material:	Not available

NMs' zeta potential values in water (pH 7 and theoretical ionic strength is  $10^{-6.998}$  mol/L) are possible to be predicted through the Enalos Cloud Platform. A kNN model is developed and validated according to OECD guidelines for the prediction of zeta potential, and it can be used in safe-by-design approaches for the design of inorganic NMs<sup>1</sup>. The properties/descriptors used for the model development (NM type, type of coating, NM core size, metal ionic radius, absolute electronegativity ( $\chi_{\text{abs}}$ ), sum of metal electronegativity divided by the number of oxygen atoms present in a particular MeOx [ $\Sigma\chi/n_o$ ]) are in turn the required input data of the online web-service (Figure 83). Data can be provided via the online form or by uploading a CSV file. The output results include the zeta potential prediction, the identities, and distances of the closest neighbours on which the prediction is based. There is also an indication whether the predicted value falls within the model's applicability domain (reliable or unreliable prediction), along with the calculated domain value and the model's APD threshold value (calculated based on the Euclidean distances approach).

A



## MS<sup>3</sup>BD Zeta Potential Predictive Model

**Note** Please that the  $\zeta$ -potential predictive model has been developed assuming that the dispersion medium is water and as such the dispersion medium pH is 7 and the theoretical ionic strength is  $10^{-6}$ – $6.998$  mol/L. Direct usage for different conditions will result in unreliable measurements.

[User Guide](#)

Nanomaterial type	Coating	Core diameter (TEM-STEM)	Metal ionic radius	$\Sigma x/nO$	Absolute electronegativity
Ag	4% (w/v) each of polyoxyethylene	19	115	1.93	4.44
Ag <sub>2</sub> S	sulfate & PVP capped	36	115	3.86	5.36
SiO <sub>2</sub>	aminoacid	86.3	40	0.95	6.19
CeO <sub>2</sub>	HMTA	30.8	102	0.56	5.65
Au	polyethylene glycol	14.5	137	2.54	5.77

ID	Zeta potential prediction	Closest NN1	Distance from NN1	Closest NN2	Distance from NN2	Closest NN3	Distance from NN3	Closest NN4	Distance from NN4	Closest NN5
Ag	-1.931	Ag JRC NM-300K	0.016	Fe <sub>2</sub> O <sub>3</sub> _SHY-058-a (PROM)	1.290	AgPURE_15nm	1.705	Ag <sub>2</sub> S 36 nm PVP	1.707	CeO <sub>2</sub> JRC NM-212
Ag <sub>2</sub> S	-42.904	AgPURE_15nm	0.034	Ag <sub>2</sub> S 36 nm PVP	1.001	CeO <sub>2</sub> (PROM)	1.179	CeO <sub>2</sub> JRC NM-212	1.181	CeO <sub>2</sub> -PVP40 (UoB)
SiO <sub>2</sub>	-22.771	SiO <sub>2</sub> -COOH (JRC) 50-60 nm	0.0	SiO <sub>2</sub> -uncoated (JRC) 50-60 nm	0.0	SiO <sub>2</sub> -COOH (JRC) 25-30 nm	0.100	SiO <sub>2</sub> -uncoated (JRC) 25-30 nm	0.100	SiO <sub>2</sub> -COOH (JRC) < 20 nm
CeO <sub>2</sub>	5.743	CeO <sub>2</sub> (PROM)	0.0	CeO <sub>2</sub> JRC NM-212	1.001	CeO <sub>2</sub> -PVP40 (UoB)	1.001	CeO <sub>2</sub> -PVP10 (UoB)	1.002	CeO <sub>2</sub> _IKR1_ (UoB)
Au	-22.435	Au 12 nm PEG-Ome (JRC-IO166D)	0.022	Au 12 nm PEG-OH/PEG-Ome (50/50) (JRC-IO166C)	0.022	AlOOH_SHY-226-B-a (PROM)	1.000	AgPURE_15nm	1.399	Ag <sub>2</sub> S 36 nm PVP

Distance from NN5	Closest NN6	Distance from NN6	Closest NN7	Distance from NN7	Domain	APD	Prediction
2.123	CeO <sub>2</sub> -PVP40 (UoB)	2.123	CeO <sub>2</sub> -PVP10 (UoB)	2.123	0.016969	1.570952	reliable
1.181	CeO <sub>2</sub> -PVP10 (UoB)	1.182	CeO <sub>2</sub> _IKR1_ (UoB)	1.182	0.034071	1.570952	reliable
0.188	SiO <sub>2</sub> -COOH (JRC) 100 nm	0.225	SiO <sub>2</sub> -NH <sub>2</sub> (JRC) < 20 nm	0.227	0.000000	1.570952	reliable
1.002	CeO <sub>2</sub> _IKS6_ (UoB)	1.003	CeO <sub>2</sub> _IKC3 (UoB)	1.004	0.000000	1.570952	reliable
1.402	CeO <sub>2</sub> JRC NM-212	1.573	CeO <sub>2</sub> -PVP40 (UoB)	1.573	0.022332	1.570952	reliable

B

Figure 83: [A] MS<sup>3</sup>BD zeta potential predictive model web service. [B] Results table including zeta potential predictions, their reliability and the neighbours' space of the input NMs.

### Nano-lazar

NMs category:	Gold NPs
Endpoint:	Cell association in A549 cell line (log <sub>2</sub> -transformed)
Reliability:	RMSE, MAE, squared Pearson correlation coefficient ( $r^2$ )
Domain:	Assessed but not provided within the tool
Availability:	<a href="https://nano-lazar.in-silico.ch/predict">https://nano-lazar.in-silico.ch/predict</a>
Training material:	Not available

This web service hosts three models built under the nano-lazar framework for the prediction of the cell association of surface-modified gold NPs to A549 cells. The three models are based on the structural, (physicochemical) property and biological (using serum protein interaction data) similarities respectively<sup>38</sup>. After the similarity calculations between the query NM and the dataset NMs, a local QSAR model is applied to acquire the prediction of the endpoint. The users should select first the model of interest and provide the respective input parameters (query NM properties), by filling the available

online forms. The results include the predicted  $\log_2$ -transformed cell association value for each query NM, and a list of the similar NMs on which the prediction is based (Figure 84).

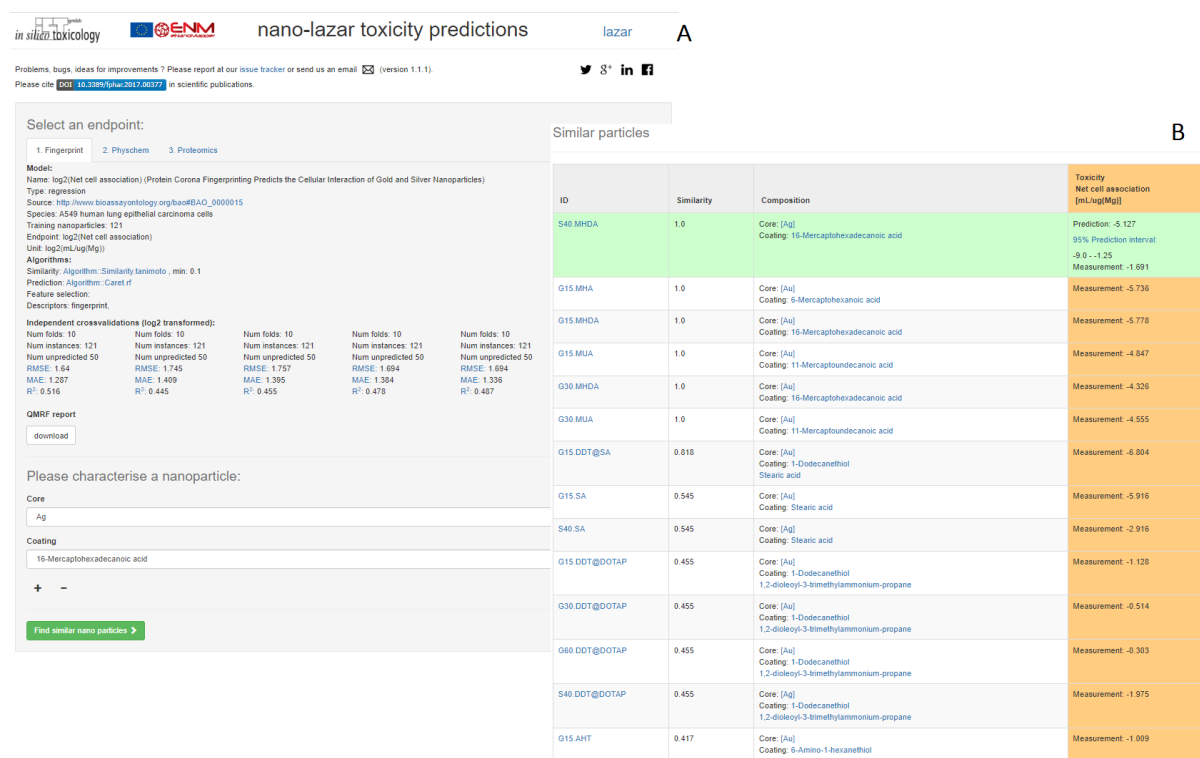


Figure 84: Screenshot of the nano-lazar web interface. [A] Selection of the model and model input parameters. [B] Results including predicted endpoint value and similar NMs to the query NM.

## Nano Protein Corona model

NMs category:	Gold NPs
Endpoint:	Cell association in A549 cell line ( $\log_2$ -transformed)
Reliability:	Coefficient of determination ( $R^2$ ), external explained variance, Golbraikh & Tropsha's test
Domain:	Indication along with the prediction
Availability:	<a href="http://www.enalosccloud.novamechanics.com/nanocommons/NanoProteinCorona/">http://www.enalosccloud.novamechanics.com/nanocommons/NanoProteinCorona/</a>
Training material:	Not available

A predictive QNAR model of the biological response of surface-modified gold NPs based on their acquired protein corona fingerprint and their physicochemical properties is created and released as a web service, to address the needs of fast-alternative screening methods for the safe-by-design of novel NPs<sup>36</sup>. The QNAR model is developed via KNIME platform using the  $k$ NN approach, and it is based on a set of significant descriptors: three physicochemical descriptors measured in serum (z-average hydrodynamic diameter, zeta potential, localised surface plasmon resonance index) and ten protein spectral counts. Similarly, in the relevant web service (Figure 85) users -for a set of untested NPs- must provide inputs for the aforementioned descriptors through the online form or by uploading a CSV file with all the necessary information. The results include the predicted  $\log_2$ -transformed cell association value for each query NP, and a warning of whether the prediction can be considered reliable based on the applicability domain of the model.

User Guide

[illegible]

Download files		
ID	Prediction (log2-transformed cell association)	Domain
"0"	-6.666127214379038	"reliable"
"1"	-7.489886474672275	"reliable"
"2"	-6.673394200613949	"reliable"
"3"	-6.569787458503942	"reliable"
"4"	-4.891621358815563	"reliable"
"5"	-5.4426938658153015	"reliable"
"6"	-6.709931735566873	"unreliable"
"7"	-6.421894327978736	"reliable"
"8"	-6.330541804250557	"reliable"

This web application was developed in the context of the NanoSolveIT project for **simulating the biodistribution of Engineered Nanomaterials (ENMs) in freshwater ecosystems**. The biokinetic model used (nanoBio) has been developed by [Gamer et al., \(2018\)](#) for describing the biodistribution of either **TiO<sub>2</sub>, ZnO or CuO** ENMs in freshwater ecosystems. The system includes **seven species**: two phytoplankton species (*S. capricornutum* and *F. crotonensis*), a zooplankton (*D. magna*), a benthic invertebrate (*H. azteca*), a bivalve (*V. constricta*), a planktivorous fish (*P. promelas*) and an upper trophic level fish (*O. mykiss*). The ecosystem also includes water, particulate matter and sediment and the ENMs can produce the corresponding dissolved metal ion via dissolution. Each species is **exposed to either the ENM or the dissolved ion** via the water, suspended particulates or dietary ingestion. Each organism is **modeled as a two-compartmental model**, one for describing the **uptake, elimination and accumulation of ENMs** and one for describing the **uptake, elimination and accumulation of dissolved metal ion**. Dissolution inside the organism is also considered by linking of the two subcompartments.

For a more details please read the corresponding publication, available at <https://doi.org/10.1021/acssuschemeng.8b01691>



This work received funding from the European Union's Horizon 2020 research and innovation program via [NanoSolveIT project](#) under grant agreement number 814572.

### Upload an Exposure Scenario

Please upload a csv file containing the exposure scenario



### Set the Simulation Parameters

Please provide the necessary simulation information

Simulation start (days)

0

Time increment (days)

1

Simulation end (days)

60

Density (kg/m<sup>3</sup>)

1280

ENM Selection

CuO

### Visualise the exposure scenario

Set the X and Y axis of the plot

X Axis

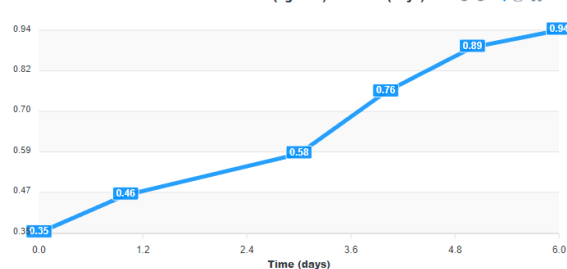
Time (days)

Y Axis

Water concentration (kg...)

☐ All TimeSeries

Water concentration (kg/m<sup>3</sup>) VS Time (days)



### Run the Simulation

### Bioaccumulation Plots

Please select the species you want to plot

☒ Plot all species

☐ F.crotonensis\_dis

☐ P.promelas

☐ D.magna\_dis

☒ O.mykiss

☐ O.mykiss\_dis

☐ S.capricornutum\_dis

☐ H.azteca\_dis

☒ H.azteca

☒ S.capricornutum

☐ D.magna

☐ V.constricta\_dis

☒ F.crotonensis

☐ P.promelas\_dis

☒ V.constricta

Predictions Plot

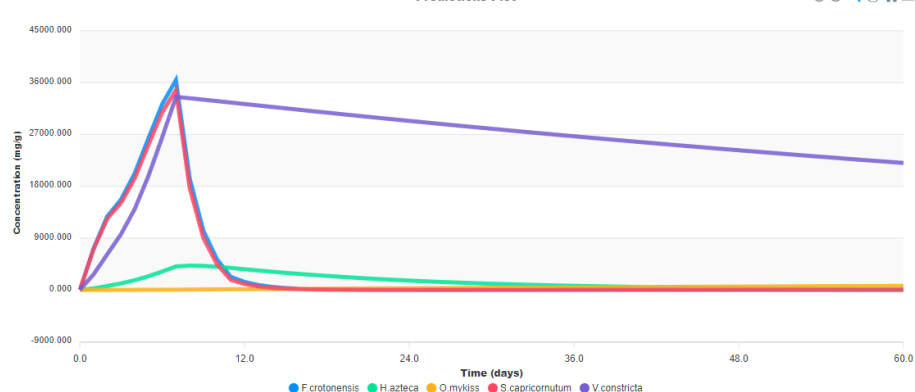


Figure 86: The NanoSolveIT NanoBio interface in Jaqpot. Users provide their data by uploading a CSV file with the exposure scenario and type the simulation parameters. The simulated NM bioaccumulation in different species time profile is presented at the bottom of the interface.

## Nanoinformatics model for zeta potential prediction powered by Enalos Cloud Platform

NMs category:	Metal and MeOx NMs
Endpoint:	Zeta potential in water (pH 6.5 or 7)
Reliability:	Pearson correlation coefficient ( $R^2$ ), external explained variance, Golbraikh & Tropsha's test
Domain:	Indication along with the prediction
Availability:	<a href="http://enaloscloud.novamechanics.com/EnalosWebApps/ZetaPotential/">http://enaloscloud.novamechanics.com/EnalosWebApps/ZetaPotential/</a>
Training material:	Written tutorial - <a href="http://www.enaloscloud.novamechanics.com/EnalosWebApps/ZetaPotential/instructions.zul">http://www.enaloscloud.novamechanics.com/EnalosWebApps/ZetaPotential/instructions.zul</a> YouTube video - <a href="https://youtu.be/swPtFTm11AI">https://youtu.be/swPtFTm11AI</a>

Zeta potential is one significant NM property which influences the NMs tendency to form agglomerates and to interact with cell membranes. A fully validated  $k$ NN/read-across predictive model is developed for the prediction of zeta potential, based on image descriptors which encode geometrical characteristics and can be easily retrieved from NMs TEM images (e.g., using an image analysis tool such as Enalos NanoXtract or ImageJ)<sup>89</sup>. Through a web service, the user can insert the indicated properties (NM's core type, experimental solutions' pH and the NMs main elongation value), which can be done either by filling in the given form or by uploading a CSV file (Figure 87A). The produced results include the predicted zeta-potential values for each NM entry as well as a warning on the prediction reliability according to the DoA limits (Figure 87B). The neighbours of each query NM are also available through the web-service, to observe any similarity patterns.

### A Nanoinformatics Model for Zeta Potential Prediction Powered by Enalos Cloud Platform

[User Guide](#)

Row ID	Type of core	pH	Main Elongation
1	oxide	6.5	
2	oxide	6.5	
3	oxide	6.5	
4	oxide	6.5	
5	oxide	6.5	
6	oxide	6.5	
7	oxide	6.5	
8	oxide	6.5	
9	oxide	6.5	
10	oxide	6.5	
11	oxide	6.5	
12	oxide	6.5	
13	oxide	6.5	
14	oxide	6.5	
15	oxide	6.5	
16	oxide	6.5	
17	oxide	6.5	
18	oxide		
19	oxide		
20	oxide		

[Download files](#)

[Execute computations](#)

B

Row ID	Predicted zeta potential [mV]	Reliability
"IO166E Gold 12 nm PEG-OH"	-0.5815616505801628	"reliable"
"AgNPs PEG 10nm"	-23.554501736363807	"reliable"
"AgNPs PEG 20nm"	-27.79291401386539	"unreliable"
"AgNPs PVP 7nm"	-29.631770892835036	"reliable"
"TiO2 103"	23.107595387224467	"reliable"
"prom ZnO"	18.615889644766884	"reliable"
"Ce NM 211"	38.20268062475627	"reliable"
"S40"	-28.832095935920286	"reliable"
"IO003"	-10.691819408446678	"unreliable"

Figure 87: [A] Screenshot of the nanoinformatics zeta potential model web interface. [B] Results table for each input NM including zeta potential predictions and their reliability. The neighbours' space is also provided by downloading the results files.

## Nanolnhale

NMs category:	TiO <sub>2</sub> NMs
Endpoint:	Biodistribution of TiO <sub>2</sub> NMs in humans after inhalation
Reliability:	Comparison to the gold standard in the calculation of particle deposition in the respiratory system (multiple-path particle dosimetry (MPPD) model)
Domain:	Users are advised to simulate scenarios for TiO <sub>2</sub> NMs of diameter close to 22 nm
Availability:	<a href="https://nanoinhale.cloud.nanosolveit.eu/">https://nanoinhale.cloud.nanosolveit.eu/</a>
Training material:	Written tutorial – available through the app

The Nanolnhale web application was developed for the simulation of TiO<sub>2</sub> NMs biodistribution in humans after exposure via the inhalation route (Figure 88). The background PBPK model<sup>161</sup> was developed based on the experimental data of Kreyling *et al.*<sup>248</sup> for TiO<sub>2</sub> NMs of 22 nm diameter and was subsequently extrapolated to humans. The application consists of two steps: the exposure scenario and the simulation information step. Users are requested to upload the TiO<sub>2</sub> concentration (in µg/m<sup>3</sup>) over time via a CSV file and to define the simulation parameters (the subject's weight, the time increment, the last time point of the simulation time vector and the size of the TiO<sub>2</sub> NMs). After data input, the application presents in graphical format the simulated internal biodistribution over time stemming out of an occupational exposure scenario. Users can select in the corresponding visualisation box which organs/tissues to plot (liver, spleen, kidneys, heart, brain, blood, Rob (Rest of the body; skeleton and remaining soft tissues), alveolar region, tracheobronchial region, lung interstitium and capillaries, upper respiratory system and, finally, lower respiratory system). It is noted that this tool implements the same model as “PBPK models and integration with the occupational exposure model” tool.

This web application was developed in the context of the NanoSolveIT project for **simulating the biodistribution of TiO<sub>2</sub> nanoparticles in humans** after inhalation exposure. The PBPK model was initially developed on rats using the biokinetic data of [Kreyling et al. \(2019\)](#) that concern **TiO<sub>2</sub> nanoparticles of 22 nm diameter** and was subsequently extrapolated to humans. Therefore, additional uncertainty should be considered in the case of distant nanoparticle sizes.



This work received funding from the European Union's Horizon 2020 research and innovation program via [NanoSolveIT project](#) under grant agreement number 814572.

### Upload an Exposure Scenario

Please upload a csv file containing the exposure scenario



### Set the Simulation Parameters

Please provide the necessary simulation information

Simulation start (hours)

0

Time increment (hours)

1

Simulation end (hours)

500

Weight (kg)

75

Size (nm)

40

Run the Simulation

### Biodistribution Plots

Please select the organs/tissues you want to plot

☒ Plot all organs/tissues

- ☐ Skin
- ☐ Blood
- ☐ Liver
- ☒ Alveolar
- ☐ Brain
- ☒ Upper.Respiratory
- ☐ Spleen
- ☒ Lower.Respiratory
- ☐ Kidneys
- ☐ Rob
- ☒ Tracheobronchial
- ☒ Heart
- ☐ Lung.Interstitium.and.Capillaries

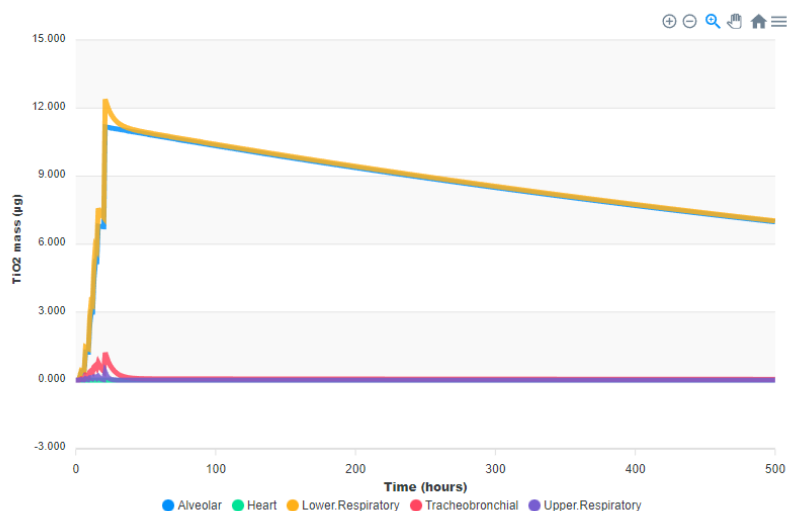
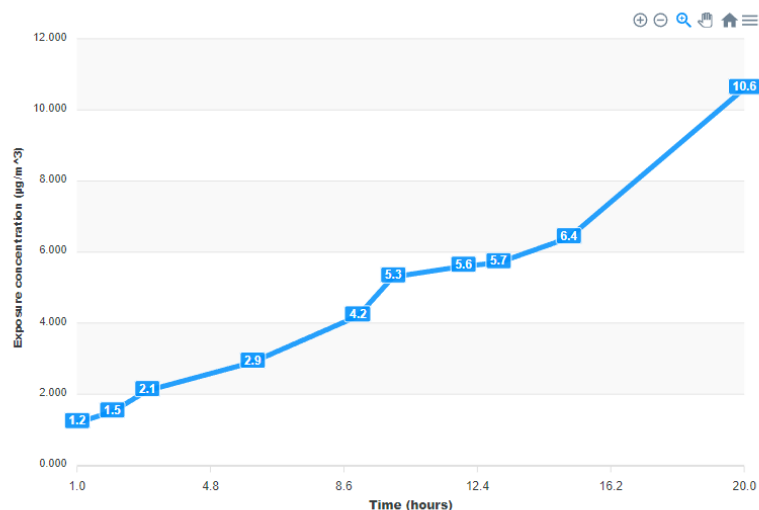


Figure 88: The NanoSolveIT NanoInhale interface in Jupyter. Users provide their data by uploading a CSV file with the exposure scenario and type the simulation parameters. The simulated TiO<sub>2</sub> concentration in different organs time profile is presented at the bottom of the interface.



## NanoMixHamster

NMs category:	TiO <sub>2</sub> -based multicomponent NMs (TiO <sub>2</sub> and Au, Ag, Pt, or Pd)
Endpoint:	Cytotoxicity to CHO-K1 cell line (expressed as pEC <sub>50</sub> )
Reliability:	Statistics provided in Mikolajczyk <i>et al.</i> 2019 <sup>100</sup> : Adjusted coefficient of determination ( $R^2_{adj}$ ), RMSE, external and internal validation coefficient ( $Q^2_{Ext}$ and $Q^2_{Int}$ ), CCC, and MAE
Domain:	Indication of the domain limits
Availability:	<a href="https://nanomixhamster.cloud.nanosolveit.eu/">https://nanomixhamster.cloud.nanosolveit.eu/</a>
Training material:	Not available

The NanoMixHamster<sup>101</sup> is an application for predicting toxicity of TiO<sub>2</sub>-based multicomponent NMs against CHO-K1 cell line (expressed as pEC<sub>50</sub>) based on SAPNet methodology<sup>112</sup>. This application is hosted in the NanoSolveIT cloud platform and consists of two main steps: the definition of the metallic NM composition (TiO<sub>2</sub> and Au, Ag, Pt, or Pd) to acquire its additive electronegativity value (Figure 89A), and the toxicity prediction step based on the results obtained in previous step (Figure 89B). The summary tab (Figure 89C) includes three graphs to help visualize the data used to develop the model and its applicability domain. Finally, the dataset generator tab can be used to generate a new dataset of theoretical TiO<sub>2</sub>-based multicomponent NMs with metal clusters of silver, and their mixtures with gold, palladium, and platinum. In a next step, it is possible to generate the predictions of the additive electronegativity and the cytotoxicity values for all the generated dataset at once.

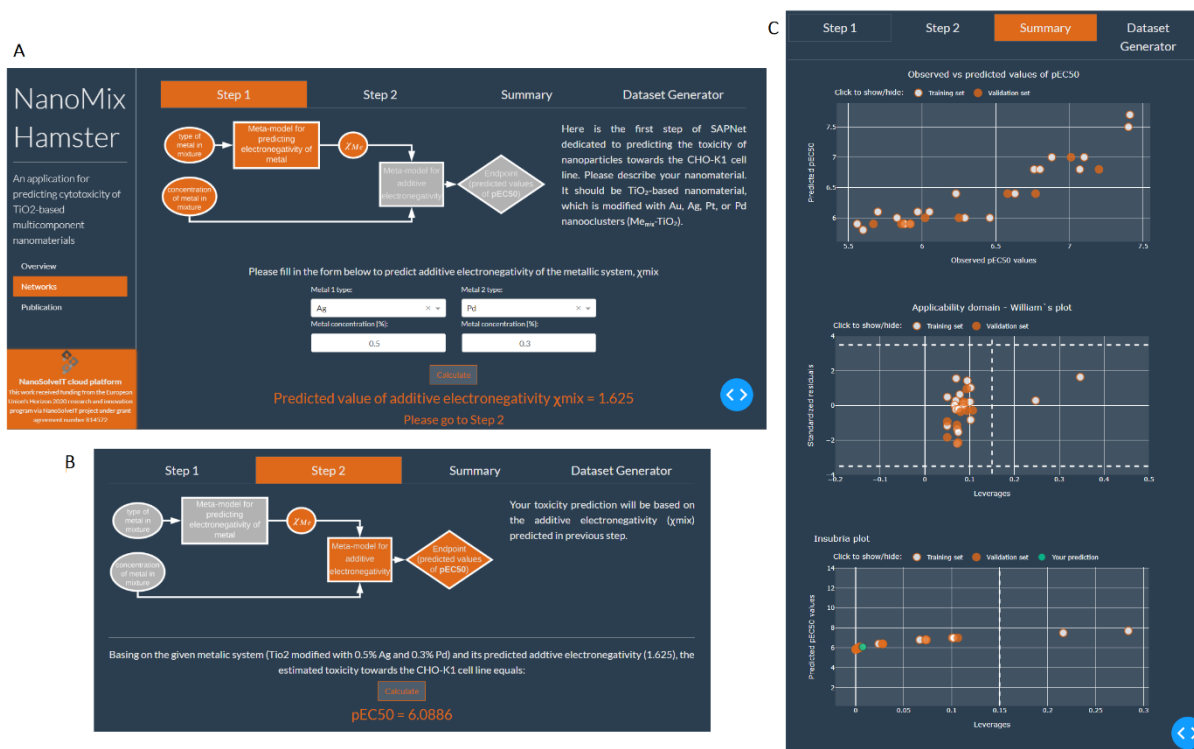


Figure 89: NanoMixHamster web interface. [A] Step 1, prediction of NM's additive electronegativity based on the metallic system composition. [B] Step 2, prediction of the NM's toxicity to the CHO-K1 cell line. [C] Observed vs predicted toxicity values and applicability domain plots.

## NanoPot

NMs category:	Au NPs
Endpoint:	Cellular uptake capacity in A549 cells, logP and zeta potential in water
Reliability:	$R^2$
Domain:	Assessed but not provided within the tool
Availability:	<a href="https://nanopot.cloud.nanosolveit.eu/">https://nanopot.cloud.nanosolveit.eu/</a>
Training material:	Written tutorial – available through the app

Nanopot is a web application hosted in Jaqpot designed to produce predictions for gold spherical NPs on three endpoints: the  $\zeta$  – potential in water, the logP ratio of concentrations of a gold NP in a mixture of octanol-saturated water and water-saturated octanol at equilibrium, and the NPs cellular uptake capacity in A549 cells. Users to initiate a prediction should provide the spherical gold NPs' PDB representation to the background regression models. The gold NP is automatically visualised in a 3D interactive scatter plot, and the prediction of the respective endpoint is also automatically displayed (Figure 90).

Figure 90: The NanoSolveIT Nanopot interface in Jaqpot. Users should select one of the three available models and upload the PDB file of a spherical gold NP of interest. After file upload, the prediction of the relevant endpoint is automatically presented on the interface.

## NanoProfiler

NMs category:	MeOx NPs and fullerenes
Endpoint:	Various (thermal conductivity, solubility, membrane damage, HIV-1 PR inhibition, smooth muscle apoptosis, bacterial reverse mutation, cytotoxicity)
Reliability:	Assessed for the models but not provided within the tool
Domain:	Assessed for the models but not provided within the tool
Availability:	<a href="https://sites.google.com/site/dtclabnp/">https://sites.google.com/site/dtclabnp/</a>
Training material:	Written tutorial available when downloading the software

This tool (Figure 91) predicts different properties of NPs using the nanoQSAR models which are included in its database and are already reported in the literature. Afterwards it performs clustering to find analogues based on the predicted property. In the tool, clustering methods are included for analogues identification, namely a *k*-Medoids algorithm (slow and exhaustive; searches best '*k*' medoids), a modified *k*-Medoid (fast; searches optimum '*k*' medoids) and a Euclidean distance-based method<sup>199</sup>.

**NanoProfiler**

Select Query Input File :

Select Output Directory :

Enter nanoQSAR Model No. (given below) :

**NanoQSARModels**

Model No.	Nano-QSAR equations
1	pEC50 (Cytotoxicity to E.coli) = 3.4056 + 0.4 (DCW(2)).
2	pEC50 (Cytotoxicity to E.coli) = 2.652 -1.342 (D3_HeteroNonMetals) + 2.014 (Sum_ActiveM_SM).
3	pEC50 (Cytotoxicity to E.coli) = 2.59 -0.5 (ΔHMe+).
4	pEC50 (Cytotoxicity to E.coli) = 4.781 -1.38 (Xox (the charge of the metal cation corresponding to a given oxide)).
5	pEC50 (Cytotoxicity to E.coli) = 4.401 -1.324 (Xox (the charge of the metal cation corresponding to a given oxide)) + 0.176 (X (m...
6	Membrane Damage (MD units 1/L) = 0.8054 + 0.1273 (DCW(2,20)).
7	Membrane Damage (MD units 1/L) = 0.7108 + 0.1336 (DCW(2,19)).
8	Membrane Damage (MD units 1/L) = 0.6128 + 0.1275 (DCW(2,21)).
9	Membrane Damage (MD units 1/L) = 10.99 -0.19 (x1(SizeInWater nm)) + 0.27 (x4(Concentration mg/L)) -1.7 (x5(ZetaPotential ...
10	Membrane Damage (MD units 1/L) = -0.3708 + 0.2378 (DCW(1,5)).

Select Clustering Method :

Number of Clusters :

Figure 91: NanoProfiler user interface.

## NanoSafer

NMs category:	Undefined - Manufactured NMs
Endpoint:	Risk Quotient (the ratio of an exposure dose to a human effect threshold)
Reliability:	Not provided
Domain:	Not provided
Availability:	<a href="http://www.nanosafer.org/">http://www.nanosafer.org/</a>
Training material:	Written tutorial - <a href="http://www.nanosafer.org/Media/Manual.pdf">http://www.nanosafer.org/Media/Manual.pdf</a>

NanoSafer is a risk assessment and control-banding tool for manufactured NMs or NP-forming processes. First, the users must log-in with their password for input data management, such as registered materials and processes. Then, for a new input material there are specific data requirements: the suppliers technical data sheet and the material safety data sheet. As for a new input process registered, required data include the type of process (powder handling or emission), information about the work area, the occupational exposure limit (OEL) for respirable dust and the energy level. Moreover, in order to make an assessment the users must choose a registered material and process, and a risk management report is generated (which can also be downloaded). Lastly, NanoSafer provides the risk level and the recommended exposure control for emission management in work scenarios based on first order modelling (Figure 92).

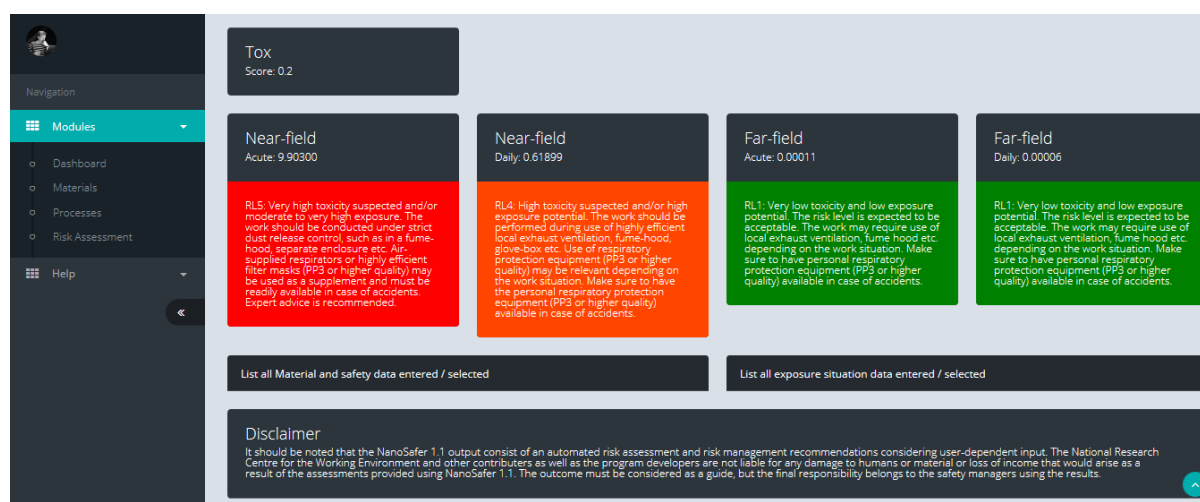


Figure 92: Screenshot of NanoSafer risk management recommendation output.

## NanoSerpA

NMs category:	Metal and MeOx NMs, graphene-based materials, fullerenes, (MW)CNTs, carbon black
Endpoint:	NMs risk assessment
Reliability:	Comparison to real scenarios reported in Literature
Domain:	Not provided
Availability:	<a href="https://www.cyc-ingenieros.com/nanoserpa/">https://www.cyc-ingenieros.com/nanoserpa/</a> for Android users only
Training material:	YouTube video - <a href="https://youtu.be/62WpDdEcnQs">https://youtu.be/62WpDdEcnQs</a>

An application for risk assessment of NMs in the insurance sector was developed<sup>200</sup>, which estimates dispersion of NMs in the environment whenever they are involved in a potentially harmful situation,

such as spills during production, transport, or use of NMs-containing products. NanoSerpA was developed by integrating hazard-related data (among others from the eNanoMapper database) and optimised exposure models while it uses probabilistic models for emission prediction, health hazard values and risk indices. This app, available for download currently only for Android smartphones, is particularly useful in the insurance sector since the user can simulate an accident scenario. Users can choose a composite material that contains NMs and provide pictures and information on the accident. Then, NanoSerpA returns a report showing the risk for the environment posed by emissions both in water, air, and soil, as well as for health, and provides a general risk index. Recommendations on the NMs use are provided according to the generated risk index. The report can be downloaded in PDF format in a standardised manner. It is also possible to access the nanoSerpA NMs database and filter the NMs according to their type and physicochemical characteristics and acquire the relevant nanotoxicity information. A terms glossary is also available through the app. The different functionalities of the app are depicted in Figure 93. Despite the usefulness of this app, it is available only in Spanish, therefore non-Spanish speakers are excluded and cannot profit from the app's functionalities.

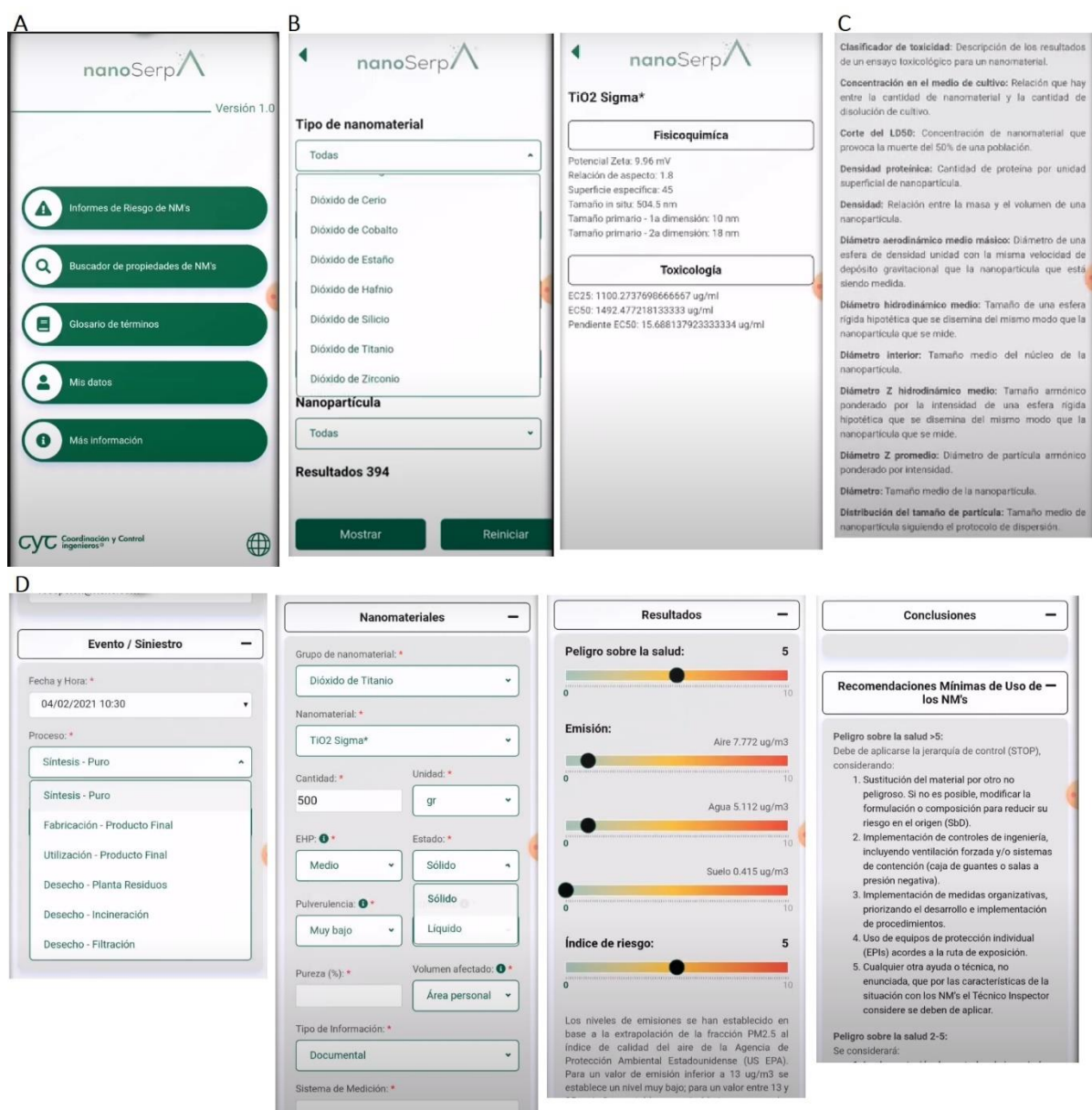


Figure 93: Screenshots from the nanoSerpA application. [A] Main page of nanoSerpA. [B] NM properties finder. [C] Terms glossary. [D] Generation of risk assessment reports, after an exposure scenario definition.

## NanoToxRadar

NMs category:	Metal and MeOx NMs
Endpoint:	Cytotoxicity probability, cytotoxicity benchmark dose, zeta potential
Reliability:	Coefficient of determination, Root Mean Squared Error, Accuracy, Specificity, Sensitivity, MCC
Domain:	Assessed but not provided within the tool
Availability:	<a href="https://nanotoxradar.kitox.re.kr/">https://nanotoxradar.kitox.re.kr/</a>
Training material:	Not available

This tool (Figure 94) allows the use of the size-dependent electron configuration fingerprint (SDEC FP) model<sup>201</sup> for the prediction of NMs cytotoxicity and zeta potential through a user-friendly interface. Users should provide the NM's core composition, the doping parameters, and coating materials from the respective dropdown menus, and the size parameters depending on the NM shape (sphere or rod). The application produces the cytotoxicity probability (more than 20% cell death of A549 cells at 100µg/mL), the cytotoxicity benchmark dose (BMD5), and the zeta potential values, along with an indication of the risk according to cytotoxicity (safe, uncertain, warning). The three background models are developed using different datasets, nonetheless a prediction can be made for all three endpoints without any indication on their reliability (based on the applicability domain of each model).

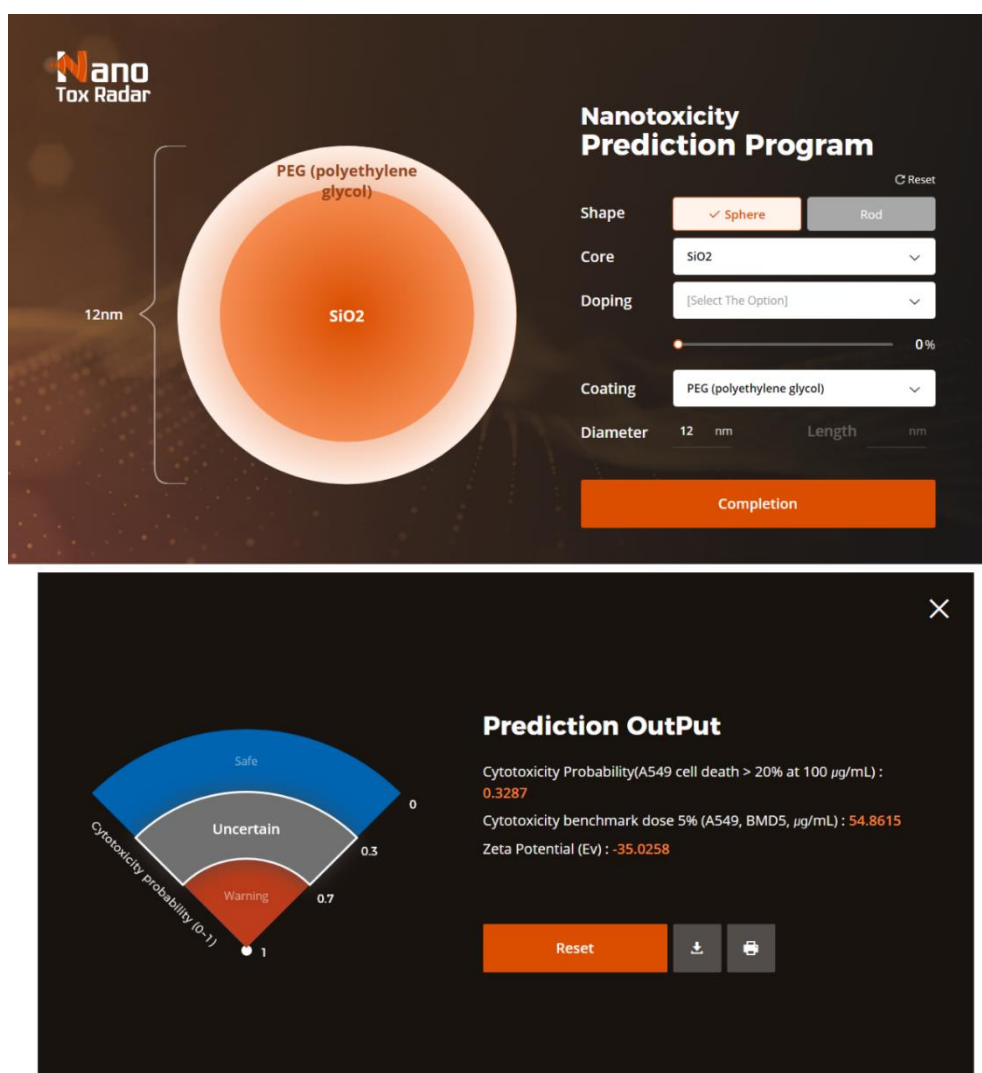


Figure 94: Screenshot of the NanoToxRadar user interface, configuration, and output windows.

## PBPK models and integration with the occupational exposure model

NMs category:	TiO <sub>2</sub> NPs
Endpoint:	NPs biodistribution in different tissues and organs
Reliability:	Comparison to the gold standard in the calculation of particle deposition in the respiratory system (multiple-path particle dosimetry (MPPD) model)
Domain:	Not provided
Availability:	<a href="https://exposurepbpk.cloud.nanosolveit.eu/">https://exposurepbpk.cloud.nanosolveit.eu/</a>
Training material:	Written tutorial - <a href="https://exposurepbpk.cloud.nanosolveit.eu/instructions.zul">https://exposurepbpk.cloud.nanosolveit.eu/instructions.zul</a>

A web application using PBPK modelling, aiding in the simulation of NP biodistribution in humans due to exposure via the inhalation route, is developed and freely available via NanoSolveIT Cloud Platform<sup>161</sup>. The application consists of two distinct modules for inhalation route exposure to TiO<sub>2</sub> NMs, an external and an internal model. In both cases the user is requested to input data for calculations initiation. External exposure scenarios (Figure 95A) can be defined directly by the user by uploading a custom-made exposure scenario according to the provided CSV template, or it can be simulated through the Multi-box aerosol model. After the definition of the external exposure model, a concentration profile as a function of time is presented. As for the internal exposure module (Figure 95B), the user must simply give information about the weight of the worker and the exposure duration (the last point of the simulation time vector in hours). The mass-time profiles are automatically generated according to the PBPK model for various body compartments (heart, spleen, brain, kidneys, alveolar, upper and lower respiratory, liver, tracheobronchial, blood, skin, lung interstitium and capillaries and, rest of the body). The user can select from the drop-down menu the desired compartment and observe the NP deposition in this organ/tissue as a function of time.

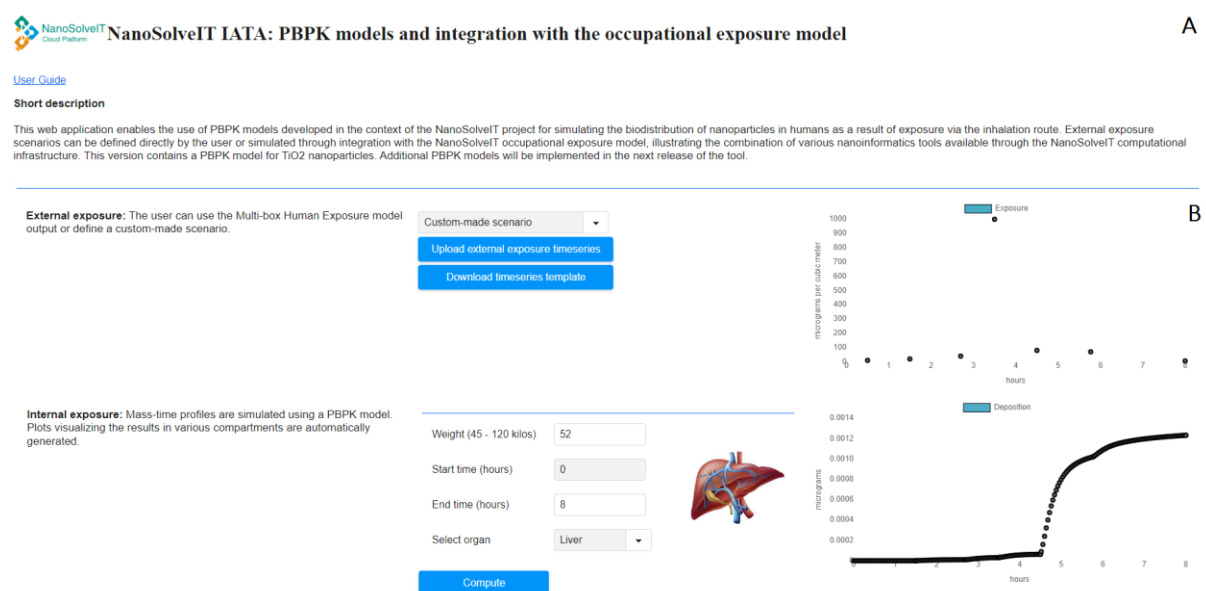


Figure 95: [A] Screenshot of the NanoSolveIT PBPK model web interface. The default timeseries template is used as a custom-made external exposure scenario [B] Simulation results on the liver following a 12-hour exposure.

## Prediction of MNPs uptake in PaCa2 cancer cells through Enalos Cloud Platform

NMs category:	MeOx NMs
Endpoint:	Uptake in PaCa2 cancer cells
Reliability:	Coefficient of determination ( $R^2$ ), external explained variance, Golbraikh & Tropsha's test
Domain:	Indication along with the prediction
Availability:	<a href="http://enaloscloud.novamechanics.com/EnalosWebApps/QNAR_PaCa2/">http://enaloscloud.novamechanics.com/EnalosWebApps/QNAR_PaCa2/</a>
Training material:	YouTube video: <a href="https://youtu.be/SKenRzr1Fbl">https://youtu.be/SKenRzr1Fbl</a>

Along with the use of manufactured NPs, concerns about potential biological effects have also increased. In this course, a fully validated PLS regression - based QNAR model is developed for the prediction of coated MeOx NPs' cellular uptake in pancreatic cancer cells (PaCa2)<sup>44</sup>. More precisely, NPs



contain a similar metal core (iron oxide/NH<sub>2</sub>) with different surface organic modifiers, thus each NP can be represented by its modifier. A user-friendly, ready-to-use online tool for toxicity predictions and NPs design is available freely via Enalos Cloud Platform (Figure 96). For prediction initiation, the user needs to provide structures for the desired compounds which can be achieved either by using the given drawing tool to design molecules/chemical structures, by entering the SMILES notations or simply by importing an SDF file. When compounds are uploaded, the web-service returns the cellular uptake value in logarithmic form ( $\log_{10}$  [nanoparticles]/cell pM) for each structure entered, accompanied by a reliability prediction according to the model's DoA. The web service provides the functionality to virtually screen a set of compounds of interest based on the validated model, and thus yielding a preliminary *in silico* testing.

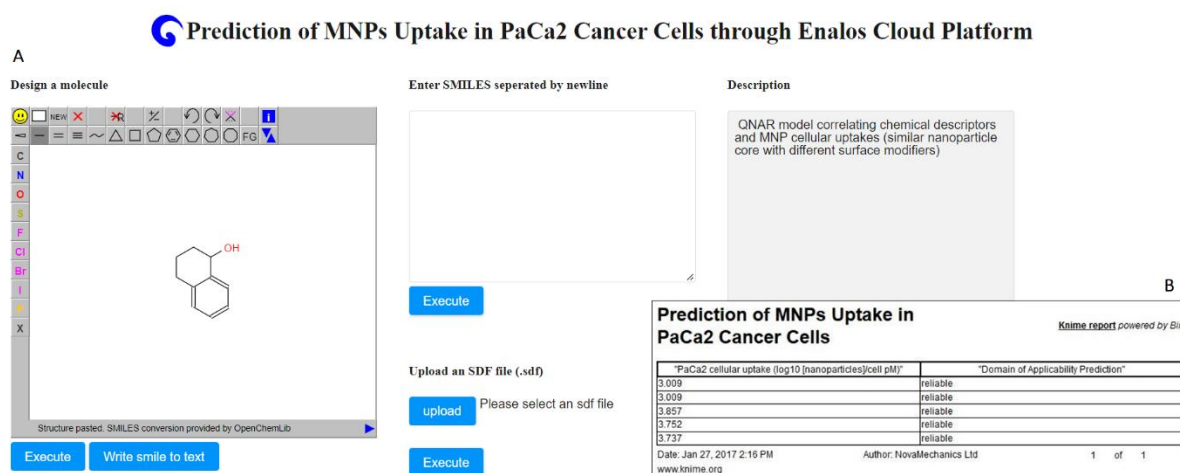


Figure 96: [A] Screenshot of the prediction of MNPs Uptake in PaCa2 Cancer Cells web interface. Different input options are available to provide the compounds (organic coatings of NPs) of interest. [B] The results table for each input compound presents the endpoint prediction and its reliability.

### QSAR for nano-mixtures

NMs category:	Binary mixtures of TiO <sub>2</sub> NPs and an organic/inorganic chemical
Endpoint:	Immobilization (%) of <i>Daphnia Magna</i> & EC <sub>50</sub> of nano-mixtures
Reliability:	Adjusted correlation coefficient ( $R_{adj}^2$ ), RMSE, MAE (for train-test and cross validation), Z score
Domain:	Assessed but not provided within the tool
Availability:	<a href="https://kriectsrc.shinyapps.io/TiO2_Dmagna/">https://kriectsrc.shinyapps.io/TiO2_Dmagna/</a>
Training material:	Not available

This user-friendly web application (Figure 97) is developed for the predictions of two ecotoxicity endpoints: EC<sub>50</sub> and immobilization (percentage of non-mobile and dead *D. magna* compared to the control sample of *D. magna*) exposed to TiO<sub>2</sub> based nano-mixtures<sup>130</sup>. The required model input data can be easily provided from the users through menus and sliders. A summary of the input data is presented after the tuning of the model parameters. Depending on the selected type of mixture descriptors, different models are employed to calculate the ecotoxicity endpoints. Models' performance is presented at the right-hand side of the interface. In both models, the predicted values for EC<sub>50</sub> or for the immobilization % percentage of the nano-mixture are automatically generated and presented along with indication of the ecotoxicity risk level (colour-coded).

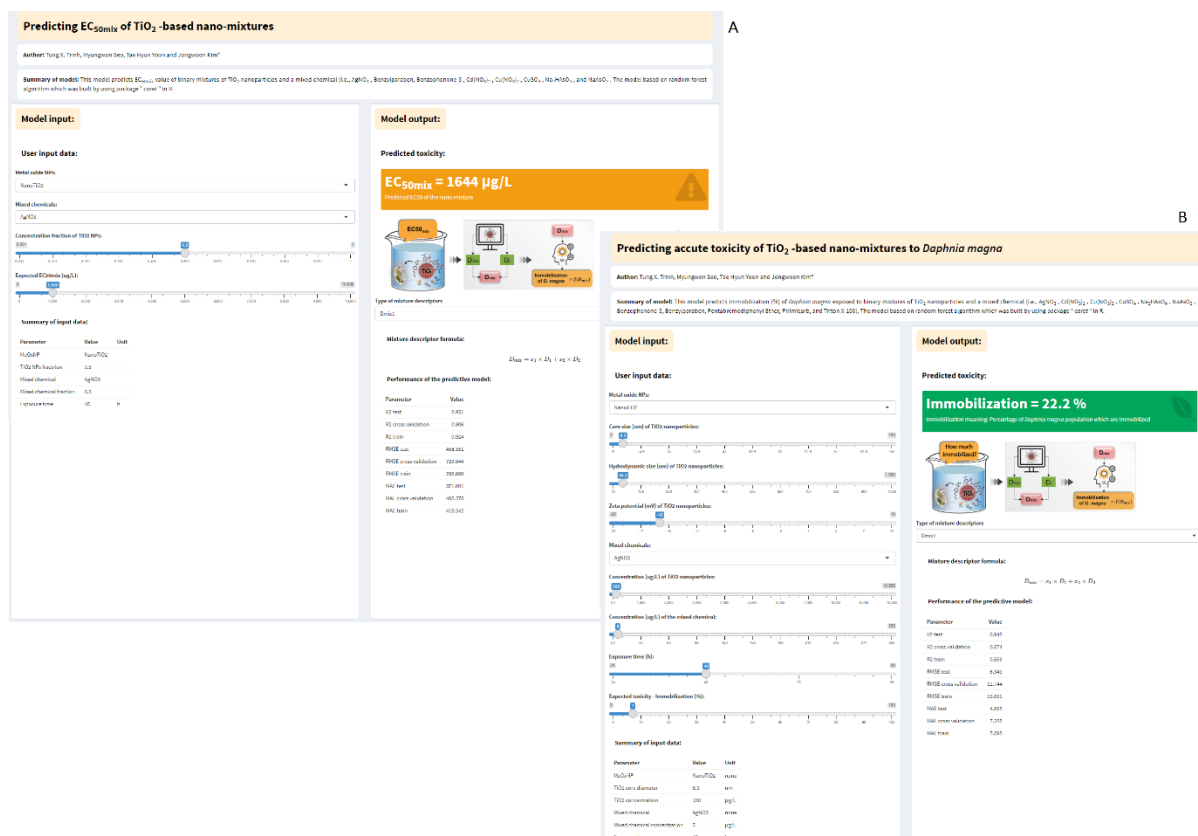


Figure 97: [A] Screenshot of the user interface of the  $EC_{50}$  of nano-mixture model. [B] Screenshot of the Immobilization (%) of *Daphnia Magna* user interface. In both cases, model input is performed at the left-hand side of the interface and predictions and an indication of level of risk (colour-based) are presented at the right-hand side of the interface.

## QsarDB models

NMs category:	MeOx NMs
Endpoint:	Cytotoxicity to bacteria <i>Escherichia coli</i> as $\log(1/EC_{50})$ and toxicity to human keratinocyte (HaCaT) cell line as $\log(1/LC_{50})$
Reliability:	Coefficient of determination ( $R^2$ ), RMSE
Domain:	Indication along with the prediction
Availability:	<a href="https://qsar.db.org/repository/handle/10967/119">https://qsar.db.org/repository/handle/10967/119</a> <a href="https://qsar.db.org/repository/handle/10967/214">https://qsar.db.org/repository/handle/10967/214</a>
Training material:	Not available

The QSAR DataBank (QsarBD, <https://qsar.db.org/>) is a platform which hosts QSAR data and models under the FAIR principles and offers integrated services that allow model analysis and use. In this course, two nanoQSAR models for the prediction of MeOx NMs toxicity-related endpoints<sup>50,63</sup> are accessible through the QsarDB. Stakeholders can access the models and for the query NM provide the indicated properties (for each descriptor, a different field is provided). Then the model is executed, and predictions are provided, along with an indication of their reliability (based on the DoA) of the models. Predictions for similar NMs based on the Euclidean distances between the query NM and training samples are also provided. Finally, the stakeholders can download the predictions and the list of similar NMs in a standardised format. In Figure 98 a screenshot of the model for the prediction of the NMs cytotoxicity to *E. coli* is presented, but a similar environment is available for the prediction of the toxicity to HaCaT cells endpoint.



chemical features. Users should provide the software with clean, ready-to-modelling data as training and test sets and tune the respective parameters including the similarity thresholds and the number of neighbours (Figure 99). Since classification-based metrics calculations are included, the tool can be used for classification read-across. The software generates the results as files and provides an evaluation of the predictions quality in terms of different validation metrics like  $q_{\text{ext},F1}^2$ ,  $q_{\text{ext},F2}^2$ , and RMSEp, if experimental data of query compounds are available. Overall, the software is impractical and not that user-friendly. Since the software is serially executed through consecutive dialog boxes, if an invalid parameter value or the wrong file is introduced, the program terminates, and all procedures must be carried out again from zero. An improved interface would allow for simultaneous display of all fields and the adjustment of the parameters in the case of a mistake or in case of repetitive execution in a sensitivity analysis scheme.



Figure 99: Data input and parameters tuning steps using the Read-Across-v4.1.

### Random forest model for the genotoxicity classification of MWCNTs

NMs category:	MWNCTs
Endpoint:	Genotoxicity (expressed via DNA strand breaks)
Reliability:	Accuracy, precision, sensitivity, specificity, MCC, F1-score, confusion matrix
Domain:	Assessed but not provided within the tool
Availability:	<a href="https://app.jaqpot.org/model/THPwkjY80z7yaIFNAYJR">https://app.jaqpot.org/model/THPwkjY80z7yaIFNAYJR</a> (upon request to become a member of the BIORIMA organization from hsarimv[AT]central[DOT]ntua[DOT]gr)
Training material:	Not available

A nanoQSAR model is developed for the genotoxicity (expressed via DNA strand breaks) prediction of MWCNTs and released as web service through the Jaqpot 5 modelling platform<sup>77</sup>. The developed random forest model classifies input MWNCTs as “genotoxic” and “non-genotoxic” (class 1 and 0 respectively), based on the values of three input features: the MWCNTs length, the % percentage of carbon purity and the hydrodynamic size of MWCNTs measured in DLS at a dose of 12.5 µg/ml (zeta average). Users after providing the values of these features to the web service (either by uploading a CSV file according to the provided template or by inserting the values in the provided form), can acquire the predicted genotoxicity values (Figure 100).

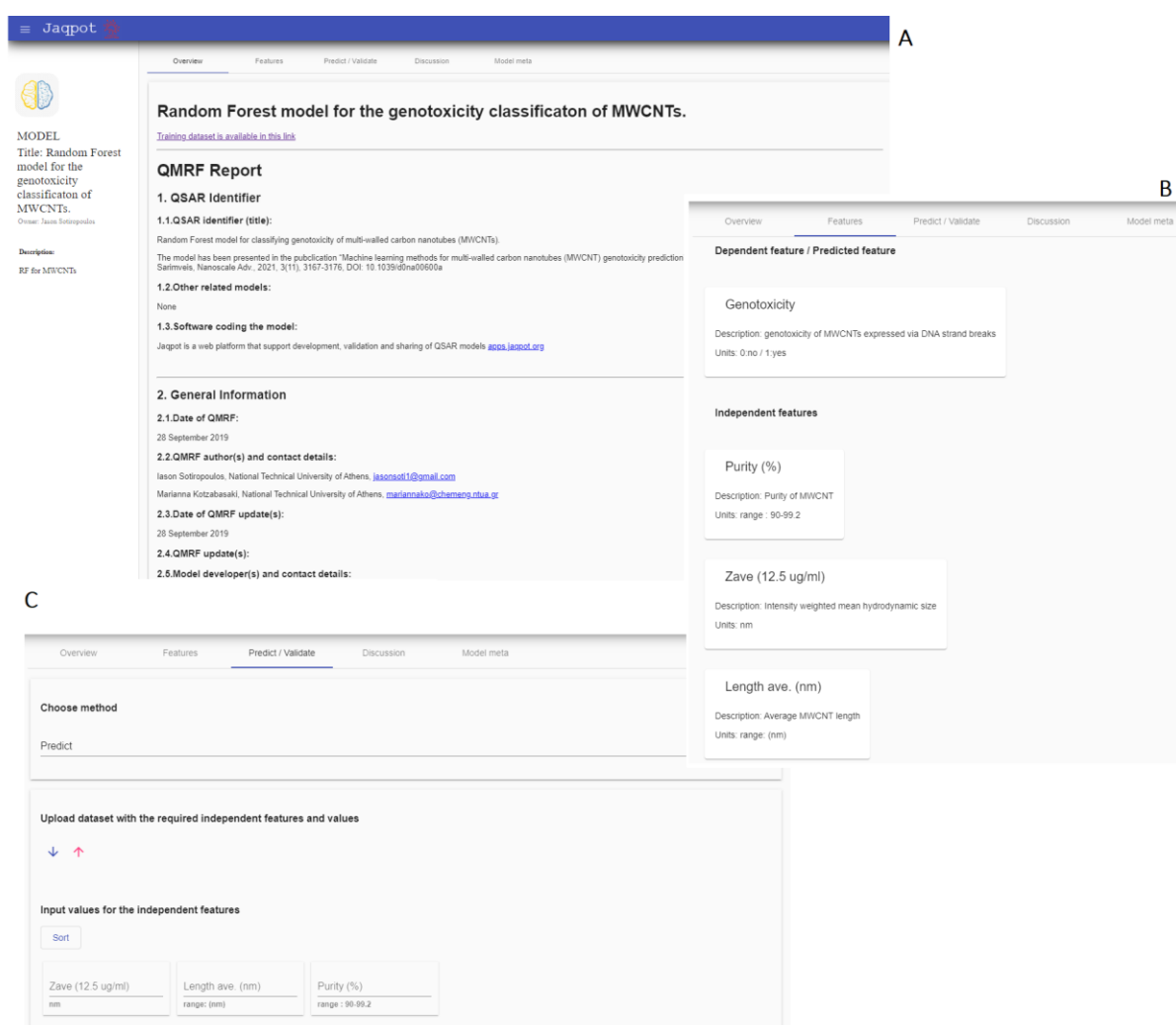



Figure 100: [A] Jaqpot random forest model for the genotoxicity classification of MWCNTs main interface. [B] Information on the necessary input variables and the predicted endpoint. [C] Prediction page where users upload their data in CSV format or via the provided online form.

### RiskGONE *in vitro* dosimetry tool

NMs category:	Metal and MeOx NPs
Endpoint:	Effective NMs concentration
Reliability:	Experimentally validated
Domain:	A cylinder space with constant dimensions of 1.5 mm diameter and 1.0 mm height
Availability:	<a href="http://www.enalosccloud.novamechanics.com/riskgone/InVitroDosimetry/">http://www.enalosccloud.novamechanics.com/riskgone/InVitroDosimetry/</a>
Training material:	Written tutorial - <a href="http://www.enalosccloud.novamechanics.com/riskgone/InVitroDosimetry/instructions.zul">http://www.enalosccloud.novamechanics.com/riskgone/InVitroDosimetry/instructions.zul</a>

A web application was developed for the numerical transport of NMs modelling<sup>203</sup>, based on a multi-step *in vitro* dosimetry methodology (developed by De Loid *et al.*<sup>249,250</sup>) For initiating the process, the user needs to select one of the nine alternative materials and one of the distribution types on the particle parameter section. In case the user wants to define a different NM, he/she must provide a density value. After that, the application requires specific solvent parameters, such as density, viscosity, and temperature. Simulation parameters must also be provided by the user; suspension column height, initial total concentration of material, centrifugation, total time of simulation in hours and time interval



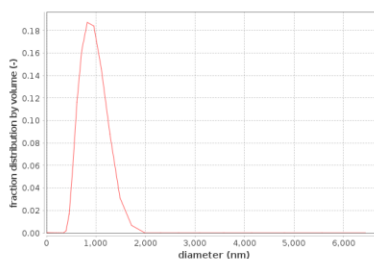
RISK GONE In vitro dosimetry

This dosimetry model is based on the Harvard DG model developed by Deloid et al

DeLoid, G.M., Cohen, J.M., Pyrgiotakis, G. et al. Advanced computational modeling for in vitro nanomaterial dosimetry. Part Fibre Toxicol 12, 32 (2015). [\[Link\]](#)

Deloid, G.M., Cohen, J.M., Pyrgiotakis, G. et al. Advanced computational modeling for in vitro nanomaterial dosimetry: Part Fibre Toxicol 12, 32 (2015). [\[Link\]](#)

Particle parameters		
Material	CeO2	▼
Density (gr/cm3)	7.215	
Distribution type	fraction distribution by volum	▼
Effective density (gr/cm3)	1.483	



Upload data ...

[Download example](#)

Solvent parameters	
Density (gr/cm3)	0.9995
Viscosity (P)	0.00081
Temperature (oC)	37

Simulation parameters	
Suspension column height (mm)	3
Height of subcompartment (mm)	0.005
Initial total concentration of material (mg/cm <sup>3</sup> )	0.1
Centrifugation (1 for gravity)	1
Total time of simulation (h)	24
Time interval for simulation (s)	0.5

**Output parameters**

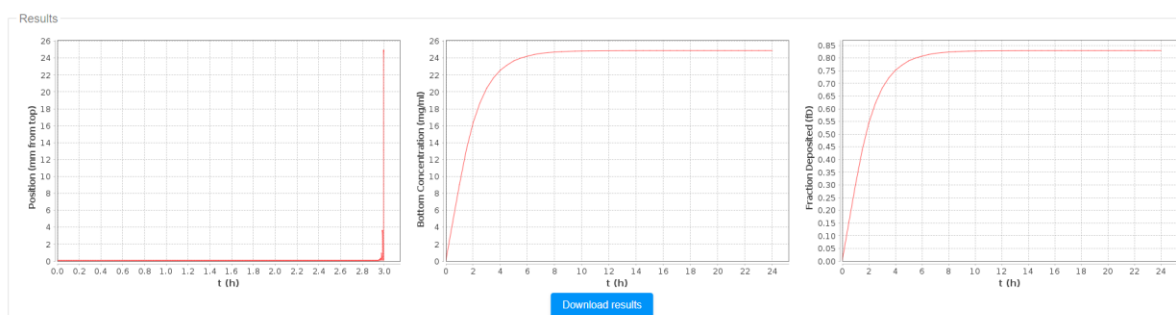
Time interval (min)

Compartment height (mm)

Write output for ... ☒ all data ☐ bottom area only

☐ Advanced parameters

Calculate

[Download results](#)

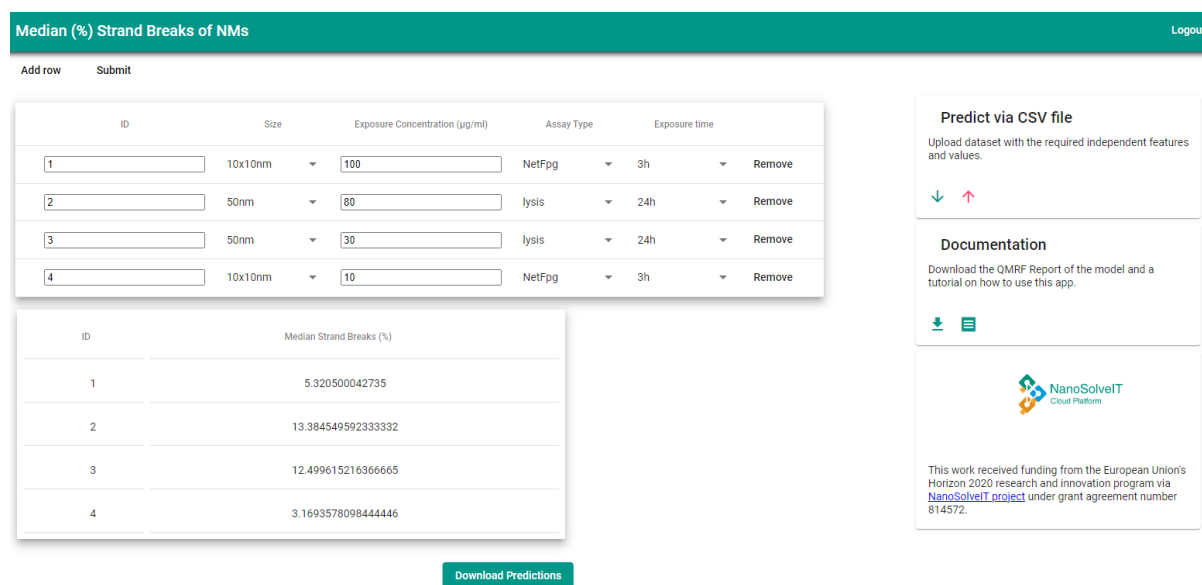
219



## Sbpot model

NMs category:	Undefined
Endpoint:	Toxicity expressed as median (%) DNA strand breaks
Reliability:	$R^2$ , MAE, MSE, RMSE, AIC, k, k'
Domain:	Assessed but not provided within the tool
Availability:	<a href="https://sbpot.cloud.nanosolveit.eu/">https://sbpot.cloud.nanosolveit.eu/</a>
Training material:	Written tutorial – available through the app

A web service that provides genotoxicity predictions is made available through Jaqpot Services, developed as part of the NanoSolveIT EU programme. A fully validated predictive model developed using the Extra Trees and AdaBoost regressors is developed, based on NMs' size, exposure conditions (concentration and duration) and the type of the Comet Assay. Users can easily access the service by creating a free Jaqpot account or via Google/GitHub accounts. Users should provide the required variable values either by completing the provided form or by uploading a CSV file following the provided template file. Results are presented shortly after the submission of the input data and include the predicted median % percentage of DNA strand breaks (Figure 102).



The screenshot displays the 'Median (%) Strand Breaks of NMs' interface. It includes a header with 'Add row' and 'Submit' buttons. The main form contains five columns: ID, Size, Exposure Concentration (µg/ml), Assay Type, and Exposure time. Four rows of data are entered. Below the form, a table shows the predicted 'Median Strand Breaks (%)' for each row. On the right, there are links for 'Predict via CSV file', 'Documentation', and a 'Download Predictions' button at the bottom.

ID	Size	Exposure Concentration (µg/ml)	Assay Type	Exposure time	Remove
1	10x10nm	100	NetFpg	3h	Remove
2	50nm	80	lysis	24h	Remove
3	50nm	30	lysis	24h	Remove
4	10x10nm	10	NetFpg	3h	Remove

ID	Median Strand Breaks (%)
1	5.320500042735
2	13.384549592333332
3	12.499615216366665
4	3.1693578098444446

Download Predictions

Figure 102: The NanoSolveIT Sbpot model interface in the Jaqpot platform. Users provide their data either manually or by uploading a CSV file. Predictions are presented at the bottom of the interface.

## Tool for assessment of human exposure to nanomaterials (multi-box aerosol model)

NMs category:	Carbon black, TiO <sub>2</sub> , and embedded TiO <sub>2</sub> NPs (TiO <sub>2</sub> –Ag, TiO <sub>2</sub> –Ag <sub>2</sub> O and TiO <sub>2</sub> –AgCl)
Endpoint:	Concentration of the NPs in an indoor environment
Reliability:	Assessed experimentally in the original publication of Jensen <i>et al.</i> 2018 <sup>251</sup>
Domain:	The computational domain is defined by the user
Availability:	<a href="https://aerosol.cloud.nanosolveit.eu/">https://aerosol.cloud.nanosolveit.eu/</a>
Training material:	Written tutorial - <a href="https://aerosol.cloud.nanosolveit.eu/instructions.zul">https://aerosol.cloud.nanosolveit.eu/instructions.zul</a>

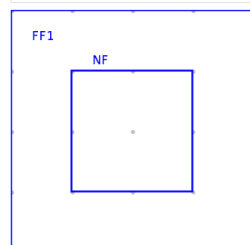
The multi-box aerosol model is an easy-to-use web application for assessment of human exposure to NPs in indoor environments. It estimates the NPs concentration in an indoor environment, e.g., a workplace laboratory or office, by dividing the space into smaller areas (near and far from the NPs emission source)<sup>161</sup>. The model is available as a free web service through the NanoSolveIT Cloud Platform (Figure 103). The prediction depends on the geometrical characteristics of the room; therefore, the computational domain must be defined by the user (geometrical layout of the room, Near Field (NF) and Far Field(s) (FF), cuboid or cylindrical). A sketch of the computational domain is provided, based on the user entries. After that, in the scenario description section, the user needs to define the NPs emission rate either by choosing a specific scenario or directly providing a distributed emission rate. Then, simulation inputs such as the modelling time, the time of active source and the number of repetitions need to be specified, while a choice to change the physical parameters (density, pressure, temperature, friction velocity) is given. The output results of the simulation consist of the evolution of the NP concentration in a particular or in all areas for different NP sizes (bins). Results can be downloaded in CSV format.

[User Guide](#)

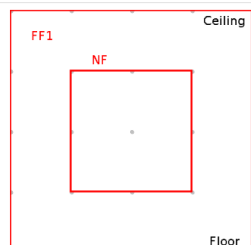
## Computational domain

Number of FF(s): 1

Area	Schema	Starting point (from lower left corner)					Schema dimensions (m)				Flow to next area (m3/min)	Flow from inlet (m3/min)	Flow to exhaust (m3/min)			
Room	Cuboid	x	0	y	0	z	0	Length	4	Width	4	Height	4	0	0	0
NF	Cuboid	x	1	y	1	z	1	Length	2	Width	2	Height	2	10	0	0
FF1	Cuboid	x	0	y	0	z	0	Length	4	Width	4	Height	4	10	0	5



ROOM TOP VIEW

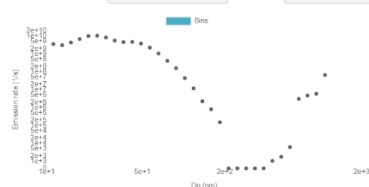


ROOM SIDE VIEW

## Scenario description

Material: Black toner Size (nm): 25.0

User defined N: 1000000000 Sigma: 2.1  $\mu$ GDP (nm): 300



## Simulation input

Modeling time (sec): 300

Time of source being active per time (sec): 50

Repetitions: 2

Include: ☒ Coagulation ☒ Deposition

Change physical parameters Density (Kg/m<sup>3</sup>): 1000 Temperature (K): 298 Pressure (atm): 1 Friction velocity (m/s): 0.1

Simulate

## Results

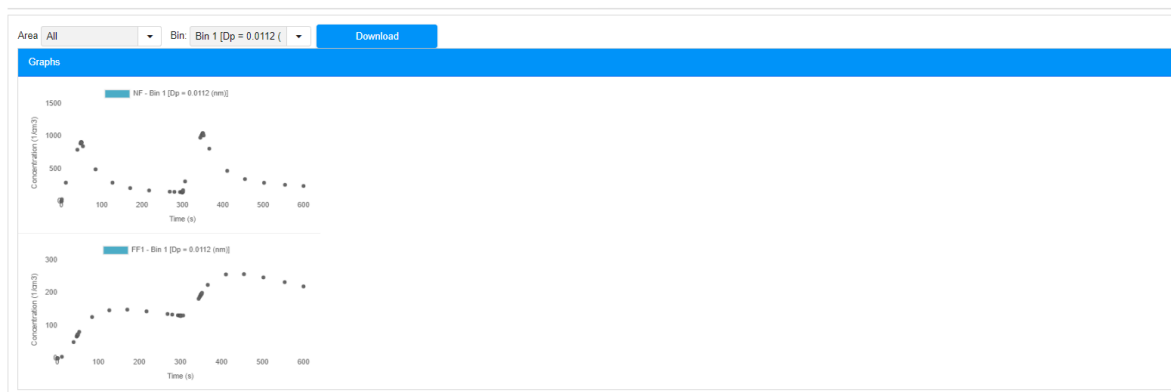


Figure 103: Interface of the NanoSolveIT multi-box aerosol model web application. In the presented scenario a 4x4x4 m<sup>3</sup> room is assumed, divided in two cubic areas: an NF and an FF of 2x2 m<sup>3</sup>. The source emits black toner particles for 50 s in a period of 300 s. The simulation is repeated twice. The user can inspect the distribution in NF and FF areas for each of the NP sizes (bins).

## toxFlow

NMs category:	General methodology for models' generation (NMs type based on input data)
Endpoint:	General methodology for models' generation (endpoint based on input data)
Reliability:	Squared correlation coefficient ( $R^2_{100}$ )
Domain:	Not provided
Availability:	<a href="https://toxflow.jagpot.org/">https://toxflow.jagpot.org/</a>
Training material:	Written tutorial - <a href="https://github.com/DemetraDanae/toxFlow/blob/master/manual%20v.Feb2020.pdf">https://github.com/DemetraDanae/toxFlow/blob/master/manual%20v.Feb2020.pdf</a> YouTube video - <a href="https://youtu.be/kGp2PuTiDrg">https://youtu.be/kGp2PuTiDrg</a>

toxFlow is a web application developed for enrichment analysis of omics data coupled with read-across toxicity prediction<sup>37</sup>. Similarity between NMs is quantified by Euclidean or Manhattan distance or by cosine similarity and the neighbours are defined considering the NMs multi-perspective characterization using one or two user-defined similarity thresholds based, for example, on physicochemical, biological, biokinetics or exposure related features. Users can train different models based on datasets of different properties and endpoints using toxFlow. In case that omics data are available it is possible to perform gene set variation analysis (Figure 104A) prior to read-across training and to include only those genes involved in statistically significant gene sets (variable selection). Data is uploaded to toxFlow in CSV format. Users can select the similarity criteria and the respective threshold values from the menus and train a read-across model. toxFlow presents the generated predictions, the accuracy of the model and the neighbours' space of each NM (Figure 104B), however the filtered variables are not provided to the users. Users can later use the produced model to acquire predictions for untested NMs through toxFlow.

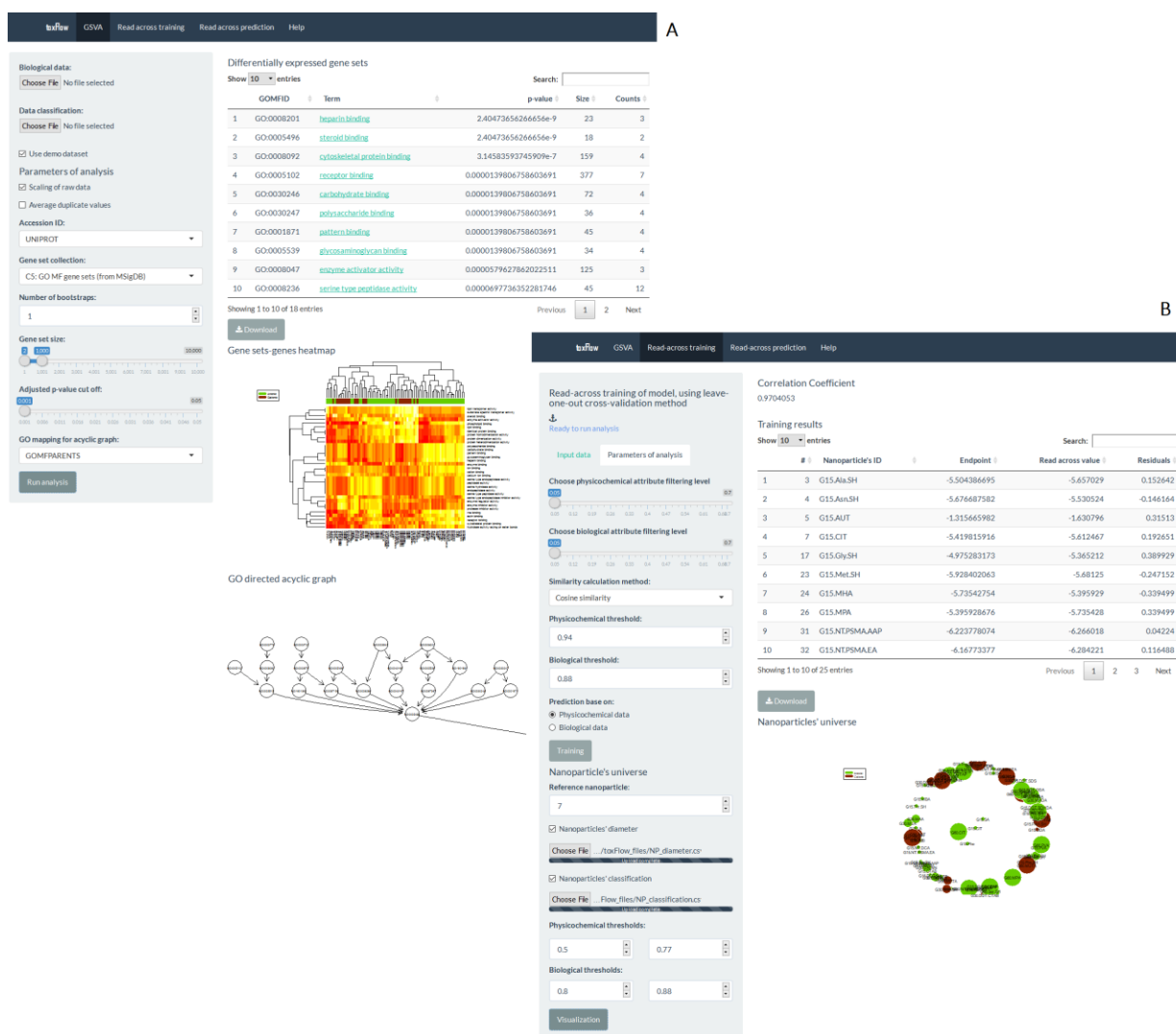


Figure 104: [A] toxFlow application web interface for gene set variation analysis. All parameters involved in the analysis can be seen in the left grey area. Results are presented in the right-hand side and include a table with all statistically significant gene sets, along with the corresponding heatmap and acyclic graph. [B] toxFlow application web interface for the training of a read-across model. All parameters involved in the analysis can be tuned from the left grey area. Results are presented in the right-hand side and include the  $R^2_{LOO}$  value of the model, the read-across prediction table with all successfully predicted NMs and the neighbours' space for a query NM.

## Vythos

NMs category:	Gold NPs & MWCNTs
Endpoint:	Cell association in A549 cell line ( $\log_2$ -transformed) & adsorption coefficient ( $\log$ -transformed)
Reliability:	Pearson correlation coefficient ( $R^2$ ), external explained variance, MSE, RMSE, MAE, explained variance in prediction (LOO)
Domain:	Indication along with the prediction
Availability:	<a href="https://vythos.jaqpot.org/">https://vythos.jaqpot.org/</a>
Training material:	Written tutorial - <a href="https://vythos.jaqpot.org/Help.pdf">https://vythos.jaqpot.org/Help.pdf</a> YouTube video - <a href="https://youtu.be/gkJuKllb_hI">https://youtu.be/gkJuKllb_hI</a>

Vythos is a web-application that hosts two predictive models developed based on datasets derived from Literature for the estimation of NM-related endpoints (toxicity or adsorption coefficients). Interested

users can upload for each of the models a set of NMs with known properties (indicated by the application) and unknown endpoint value and acquire reliable predictions. Vythos incorporates the concepts of grouping, read-across and optimization<sup>35</sup>. Based on the formulation and the solution of a mathematical MILP problem, the method used to build the models, searches over a space of alternative grouping hypotheses, in terms of partition feature (group defining property), breakpoints (group limits), and selected features in each region and determines the one providing the most accurate read-across predictions. One the incorporated models is based on the definition of groups in two dimensions (using two partition features). Predictions for untested NMs are produced from region-specific linear models depending on the group in which each query NM belongs based on its partition feature value. To produce the predictions of the endpoint of interest, users should choose the respective model and upload a CSV file with all the values of the necessary variables (indicated in the parameters field). Templates are available for each of the models. The application (Figure 105) presents for any input NM data the endpoint predictions and their reliability, as well as information about the group in which they belong (region ID and interactive graph).

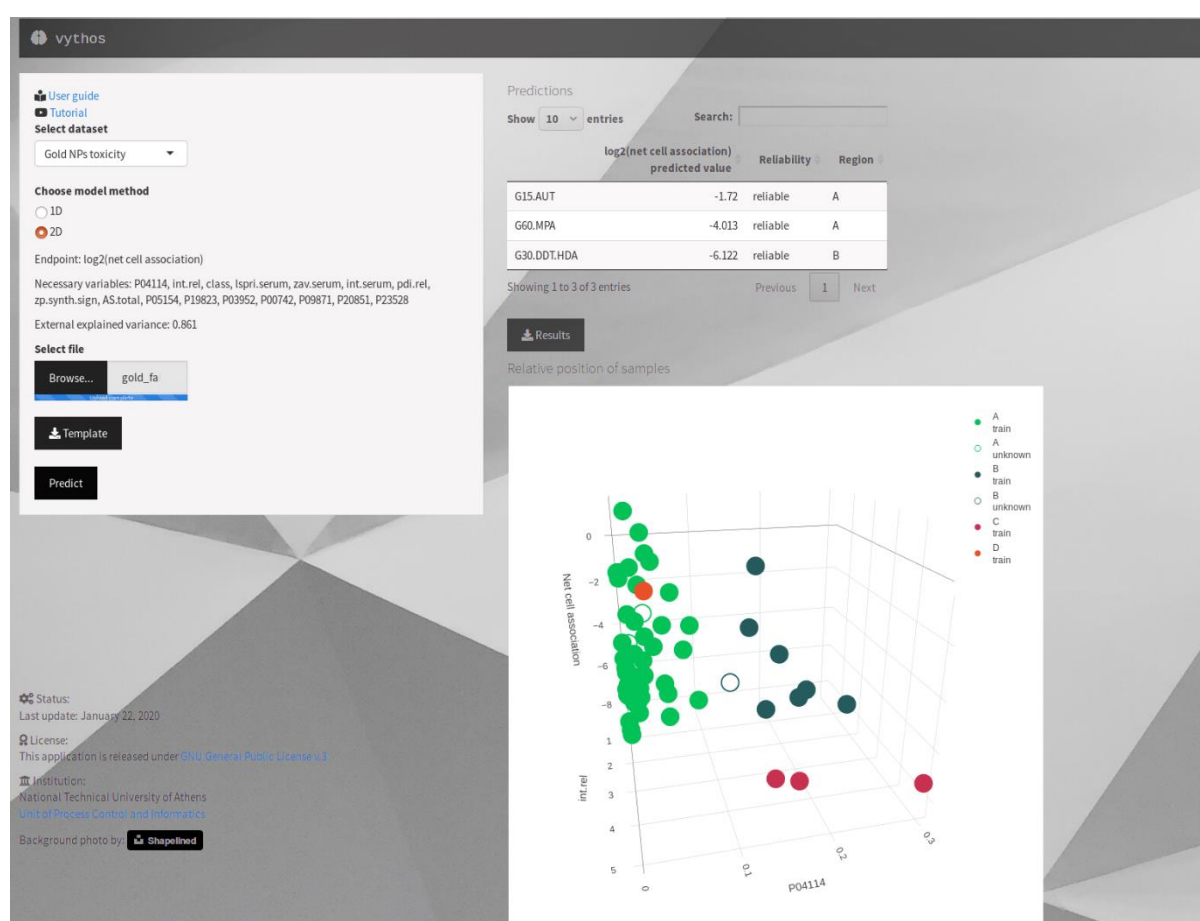


Figure 105: vythos web application interface. Users select the model they want to use and acquire the input NMs predictions, their reliability and information on the neighbours' space.

## ζ-regression

NMs category:	Metal and MeOx NMs
Endpoint:	Zeta potential in water (pH 7)
Reliability:	$R^2$
Domain:	Assessed but not provided within the tool
Availability:	<a href="https://zetapot.cloud.nanosolveit.eu/">https://zetapot.cloud.nanosolveit.eu/</a>
Training material:	Written tutorial – available through the app

Another web service that provides zeta potential predictions in water for metal and MeOx NPs is made available through Jaqpot Services, developed as part of the NanoSolveIT EU programme. A fully validated predictive model developed using the XGBoostTree algorithm based on NPs' physicochemical and molecular descriptors is developed. Users can easily access the service by creating a free Jaqpot account or via Google/GitHub accounts. Users should provide the required variable values either by completing the provided form or by uploading a CSV file following the provided template file. Results are presented shortly after the submission of the input data and include the predicted zeta potential (Figure 106).

### ζ - Regression

Add row Submit

ID	Diameter(nm)	Molecular Weight (g)	Coating Charge	Surface Area (nm <sup>2</sup> )	Electronegativity	Coating	# of Oxygen	
1	19	107.8662	Neutral	1134.114948	1.93	Yes	0	Remove
2	36	247.8	Neutral	4069.44	1.93	Yes	0	Remove
3	47.3	107.8662	Anionic	7025.0906	1.93	Yes	0	Remove
4	42.56487921	172.115	Neutral	5691.84	1.12	Yes	2	Remove
5	4.5	172.115	Anionic	63.61725124	1.12	Yes	2	Remove
6	3	172.115	Cationic	28.27433388	1.12	Yes	2	Remove
7	83	172.115	Neutral	21642.43179	1.12	Yes	2	Remove
8	8.2	172.115	Neutral	211.24069	1.12	No	2	Remove
9	7.625	172.115	Neutral	182.6541604	1.12	Yes	2	Remove
10	8.992	172.115	Neutral	254.0168167	1.12	Yes	2	Remove

#### Predict via CSV file


Upload dataset with the required independent features and values.

↓ ↑

#### Documentation

Download the QMRF Report of the model and a tutorial on how to use this app.

↓



This work received funding from the European Union's Horizon 2020 research and innovation program via [NanoSolveIT project](#) under grant agreement number 814572.

ID	Zeta (mV)
1	-7.359
2	-42.7438
3	-53.0915
4	-3.4211
5	1.1476
6	-3.1649
7	-8.1325
8	-5.2671
9	-0.6078
10	6.5846

Download Predictions

Figure 106: The NanoSolveIT ζ-regression model interface in Jaqpot. Users provide their data either manually or by uploading a CSV file. Predictions are presented at the bottom of the interface.

## Appendix III – Questionnaire on valid *in silico* modelling tools and read-across approaches

In the next paragraphs the questionnaire prepared and distributed to collect the experts' opinion is presented. Its online form is slightly modified to ensure interactivity with the experts. A short introduction (used as cover letter) to inform the experts for the scope of this survey is also presented.

### Preamble/Cover letter

Dear Colleagues,

We are conducting a survey on the existing *in silico* modelling tools and read-across approaches for nanomaterials on behalf of the EU Observatory for Nanomaterials (EUON) and the European Chemicals Agency (ECHA). The aim is to map the state-of-the-art situation in the field of computational-alternative methodologies for the hazard and risk assessment of nanomaterials, including (nano)QSAR and grouping/read-across modelling, AOP methodologies, and physics-based and PBPK/PBTK modelling. In addition, the study aims at examining what tools are available for the *in silico* assessment of nanomaterials adverse effects and properties estimation.

The key aim of this survey is to identify the data gaps or missing steps to be applicable for research and regulatory purposes and to identify the development needs to further improve this area and make them regulatory compliant and usable.

Therefore, we are sharing this questionnaire, which should not take more than 15 minutes to complete, to help us gather the necessary information. This questionnaire is structured in five sections (including the personal information section), to assess the developed methodologies, on which data are based, how are they disseminated and what are the gaps that should be addressed in the future. If you are actively involved in developing models and *in silico* methodologies for nanomaterials, you will be enabled to answer closed ended questions mapping the characteristics of these methods. We would appreciate your participation in our study.

Following participation, if you wish, you can gain early access to the report that we will compile with EUON/ECHA. In any case, any information shared with us will be treated, if desired, as confidential and will not be published.

Questionnaire link: <https://ec.europa.eu/eusurvey/runner/InSilicoReadAcrossApproaches>

If you don't wish to fill in the questionnaire and would prefer a live interview, we would be happy to arrange this with you. Please mail us at [surveys@novamechanics.com](mailto:surveys@novamechanics.com).

We would like to thank you in advance for your cooperation and active participation in completing this exercise.

Antreas Afantitis, PhD, MBA

Managing Director

e-mail: [afantitis@novamechanics.com](mailto:afantitis@novamechanics.com)

NovaMechanics Ltd

Cheminformatics & Nanoinformatics Excellence

Digeni Akrita 51, 1070, Nicosia, Cyprus

Correspondance Address: P.O Box 26014, Nicosia, 1666, Cyprus

<http://www.novamechanics.com>



## Section 1 – Personal information

- \*Name<sup>l</sup>
- \*Affiliation
- Position
- \*e-mail
- Are you part of an EU, national or international project? (Yes, No)
  - If yes, in which project are you participating?

## Section 2 – *In silico*, grouping, and read-across methodologies

- Are you aware of existing grouping and read-across frameworks?
  - If yes, can you tell us about the frameworks you are aware of?
- Have you used any *in silico* method or tool for research or regulatory purposes?
  - If yes, can you, please, explain in which context you have used them?
- Are you developing or involved in developing models and *in silico* methodologies for nanomaterials?<sup>m</sup>
  - If yes, can you, please, tell us the models and *in silico* methodologies you have developed or have been involved with developing?
  - Which type of methodologies (e.g., nanoQSARs, grouping/read-across, PBPK models, AOPs, physics-based, etc.) are you employing for the developed of *in silico* approaches?
  - Are the developed *in silico* methodologies dataset-dependent?
  - What are the standard validation methodologies you employ for the developed *in silico* methods?
  - Which are the metrics used in model validation?
  - How do you ensure the reliability of the predictions of the developed methodologies?
  - How is the assessment of the applicability domain of the developed methodologies being performed?
  - How is model interpretation performed?
  - Do you include model documentation using a specific template (e.g., QMRF, MODA) when releasing an *in silico* model?
- Are the developed *in silico* methodologies dataset-dependent?
- \*Should any provided information be treated as confidential?

## Section 3 – Data availability

- In your opinion, are there any barriers in the data availability in nanoinformatics?
  - If yes, which are the current availability barriers?
- Do you believe that datasets of sufficient quality exist for developing *in silico* methodologies?
  - If yes, is it easy to access such datasets (e.g., publicly available)?
    - If no, what are the accessibility bottlenecks?
- Are you developing or involved in developing models and *in silico* methodologies for nanomaterials?
  - If yes, which are your main sources of data?

---

<sup>l</sup> Fields marked with \* are mandatory.

<sup>m</sup> This question is repeated through the questionnaire (marked in grey colour). However, in its' online version it will be presented once. More specifically, an interactive scheme is implemented: each expert answers once in the question of "developing new *in silico* methodologies" in section 2 and the depended sub-questions (marked with different bullet-format) in sections 2,3 and 4 are presented accordingly.

- Do you include data quality control as a standard procedure prior to any modelling activities?
  - If yes, what data quality procedures do you follow?
- Do you prefer to use data from sources compliant with the FAIR and Open Data guidelines over other sources of data (FAIR/Open Data/I don't have a preference)?
- Are there any preferences in the type of properties used for modelling purposes? (e.g., physicochemical, *in vitro*, *in vivo*, atomistic, molecular, period table, image properties, etc.)
- In your opinion, should data collation from different sources be performed in case of data scarcity?
- What are the data collation barriers?
- \*Should any provided information be treated as confidential?

## Section 4 – Dissemination

- In your opinion is the dissemination of the developed *in silico* methodologies as user-friendly workflows necessary to support the regulatory applications and the SbD of novel nanomaterials?
- Are you developing or involved in developing models and *in silico* methodologies for nanomaterials?
  - If yes, how do you make your methodologies available to potential stakeholders (e.g., as web-tool, stand-alone software, etc.)?
  - Is there a licensing system for using your tools?
  - Are you models freely available or is there a fee for using? (Free/Freeware/One-time free/Subscription)
  - Are you providing the source code of the tools you are developing?
  - Do you provide APIs for remote accessing of the *in silico* tools?
- Are you aware of any *in silico* methodologies integrated in industrial, regulatory or research level for the SbD, grouping, and read-across of nanomaterials?
  - If yes, which *in silico* methodologies do you know that are integrated in industrial, regulatory or research level for the SbD, grouping, and read-across of nanomaterials?
- What is the confidence of relevant stakeholders regarding the reliability of the developed models and *in silico* methodologies?
- \*Should any provided information be treated as confidential?

## Section 5 – Gaps and future steps

- In your opinion, which are the current gaps in the *in silico* investigation of nanomaterials?
- Is it possible to integrate the results of the *in silico* methodologies in the SbD of novel nanomaterials or regulatory applications?
- Can *in silico* methodologies used for bulk chemicals be used in the case of nanomaterials? (Yes/No/With modifications)
  - If yes, which bulk chemical methodologies can be used for nanomaterials?
  - If “with modifications”, which bulk chemical methodologies can be used for nanomaterials? What are the required modifications?
- What are the difficulties in the integration of the results of the *in silico* methodologies in the SbD of novel nanomaterials or regulatory applications?
- How do you think it would be possible to reinforce the confidence of the stakeholders in reliability of modelled results and *in silico* methodologies?
- Which are the key parameters future *in silico* investigation of nanomaterials should focus (on)?

- Would you like to share any other thought/comment regarding the development and use of *in silico* methodologies for nanomaterials?
- Would you be interested to give us a personal perspective regarding *in silico* method development and application of nanomaterials?
  - If yes, can we contact you for further discussion?
- \*Should any provided information be treated as confidential?

## Appendix IV – Graphs with the responses to the questionnaire

The responses from the closed questions of the questionnaire are presented in graphical format.

### Section 1 – Personal information

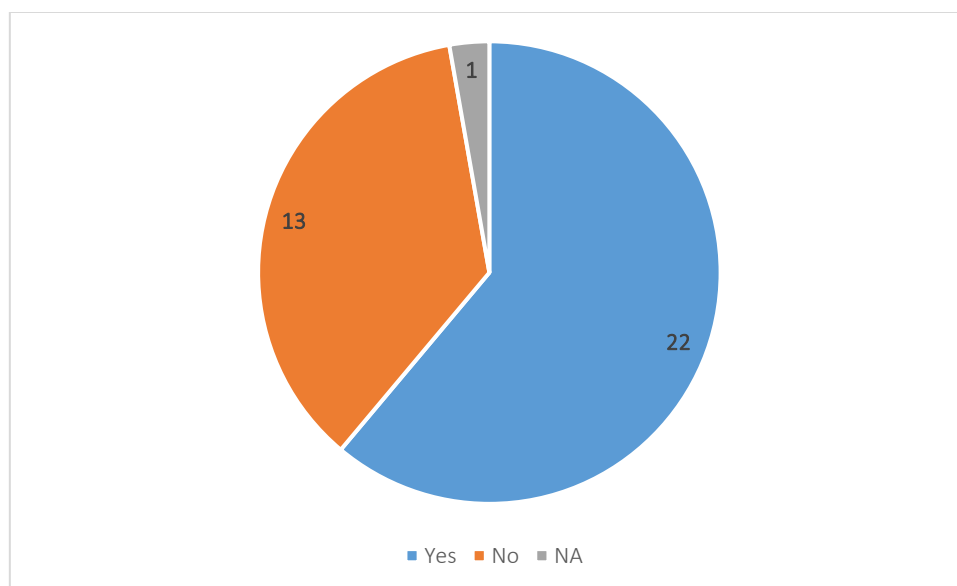


Figure 107: Are you part of an EU, national, or international research project?

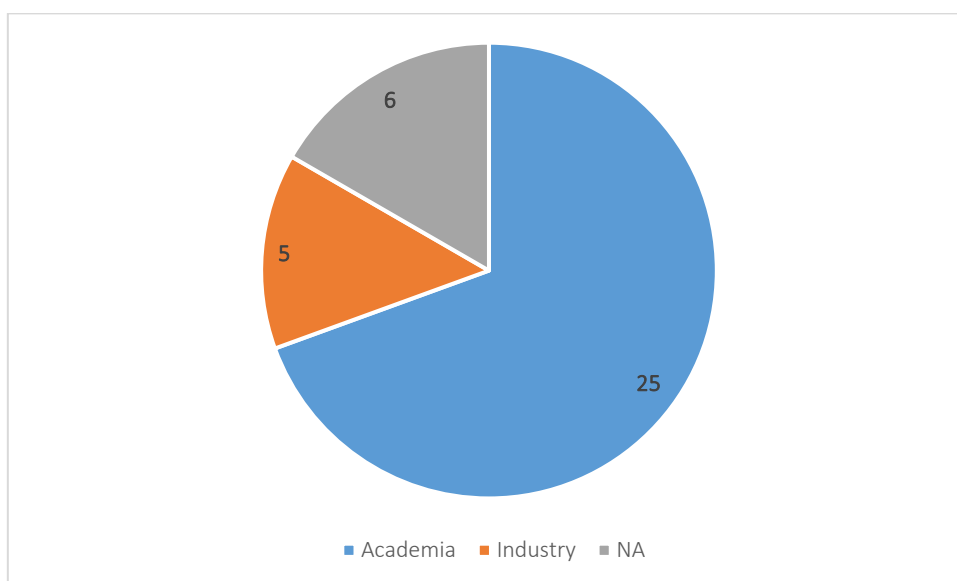


Figure 108: Experts' affiliated sector.

## Section 2 – *In silico*, grouping, and read-across methodologies

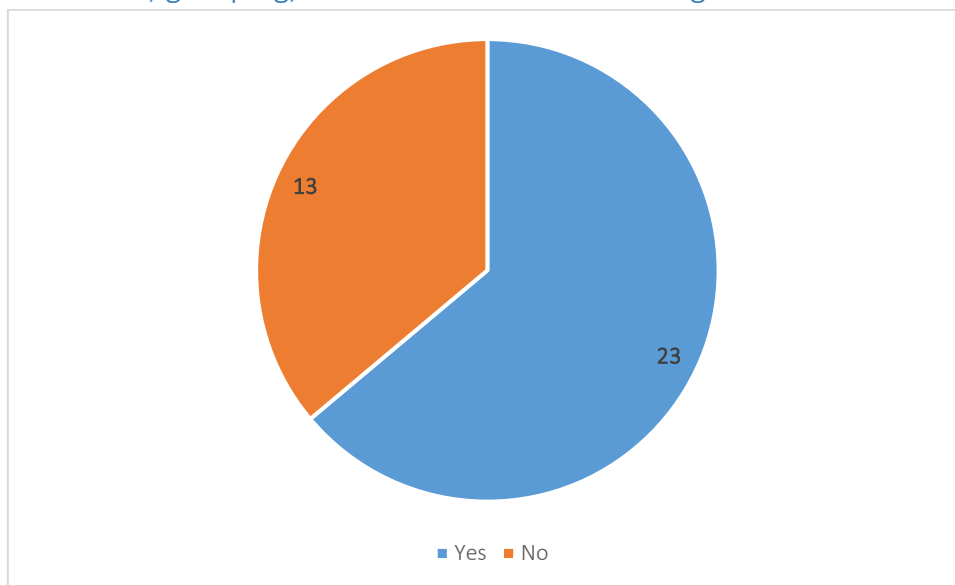


Figure 109: Are you aware of existing grouping and read-across frameworks?

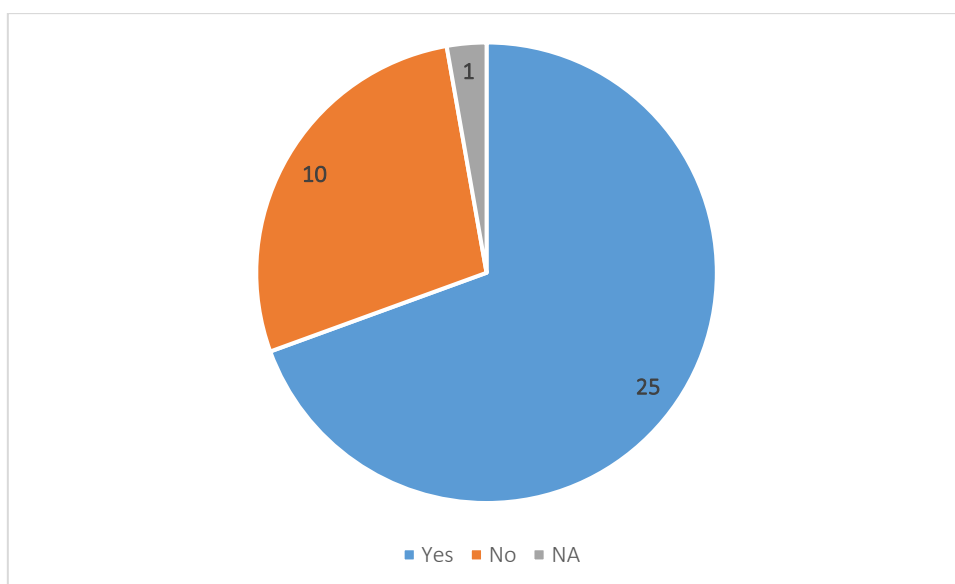


Figure 110: Have you used any *in silico* method or tool for research or regulatory purposes?

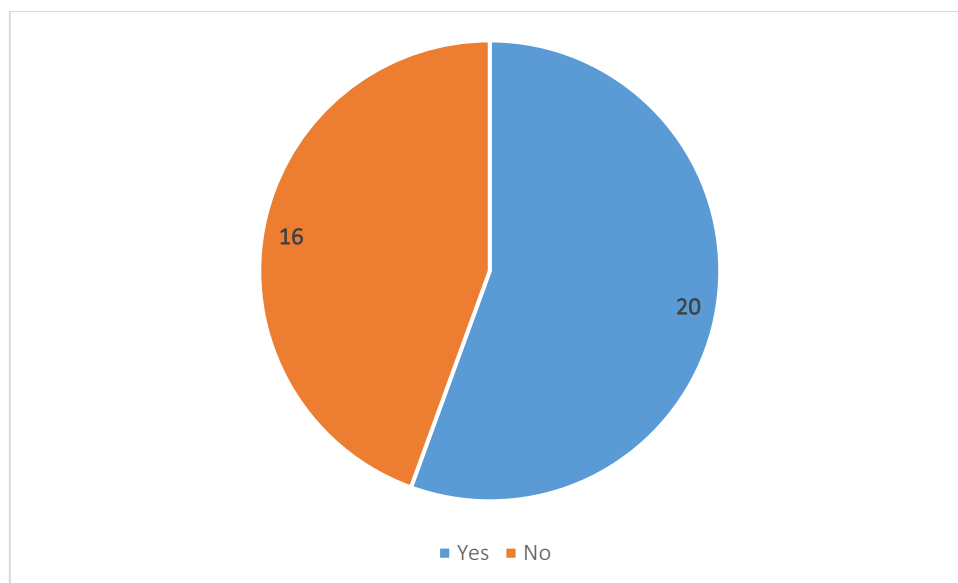


Figure 111: Are you developing or involved in developing models and *in silico* methodologies for nanomaterials?

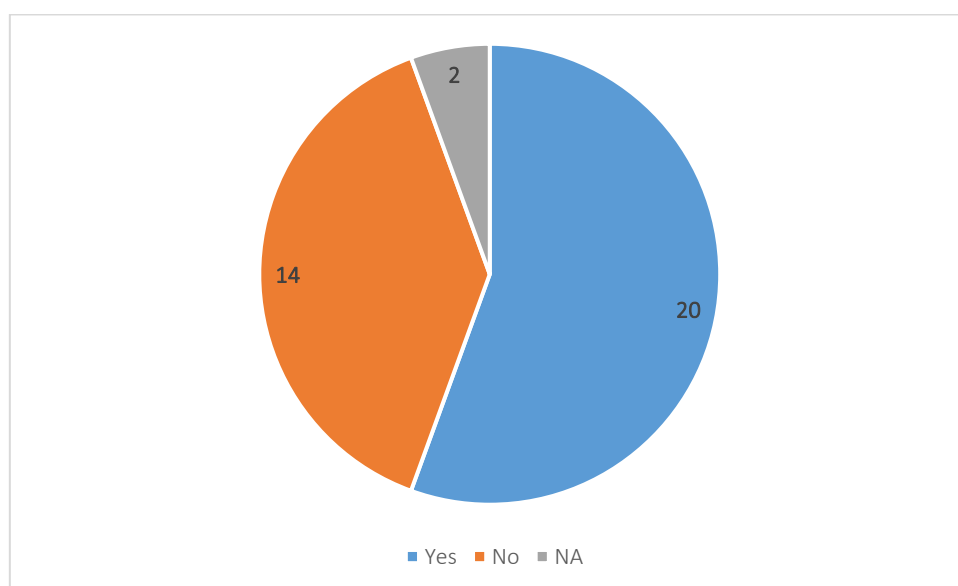


Figure 112: Are the developed *in silico* methodologies dataset-dependent?

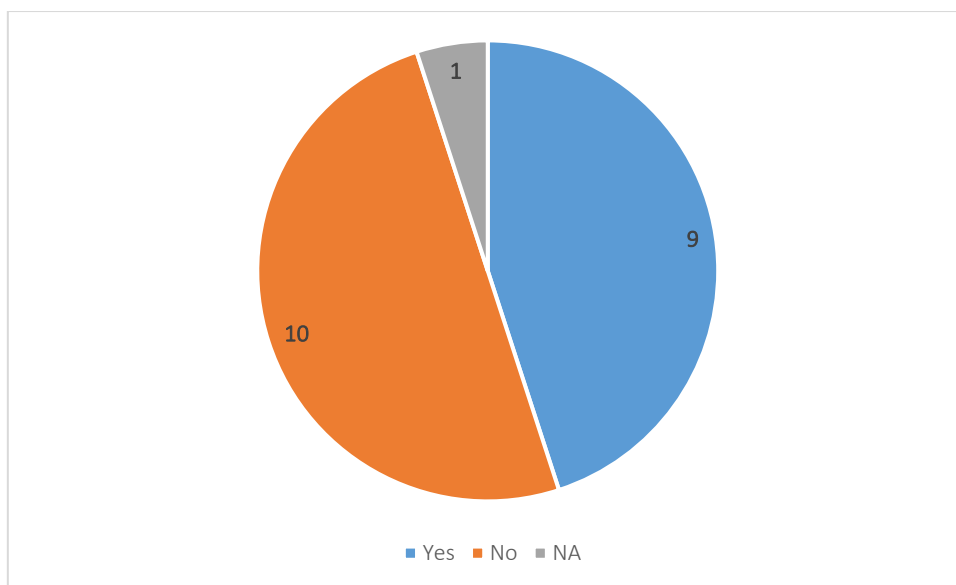


Figure 113: †Do you include model documentation using a specific template (e.g., QMRF, MODA) when releasing an *in silico* model? Graphs starting with “†” symbol present only the answers of the experts who are developing or involved in developing models and *in silico* methodologies for NMs.

### Section 3 – Data availability

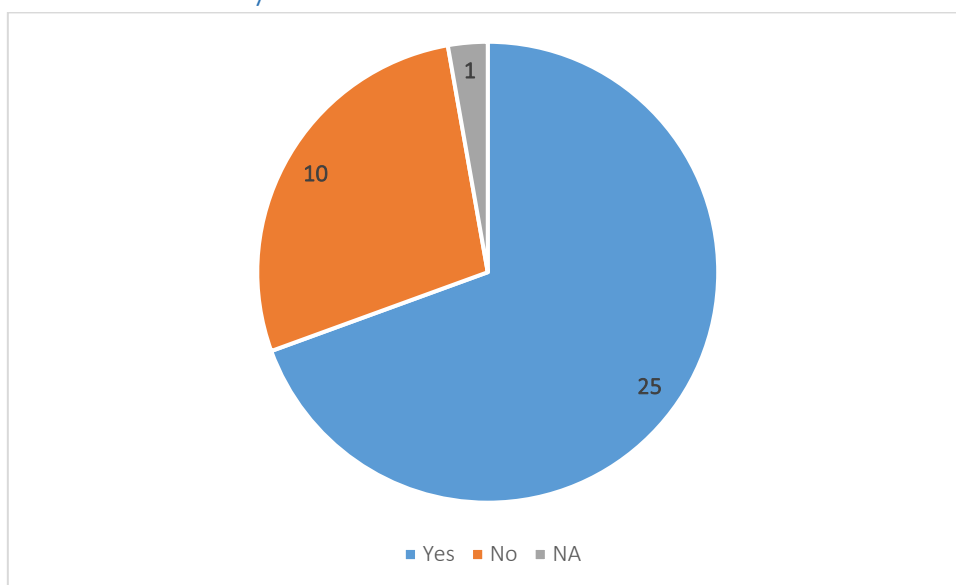


Figure 114: In your opinion, are there any barriers in the data availability in nanoinformatics?

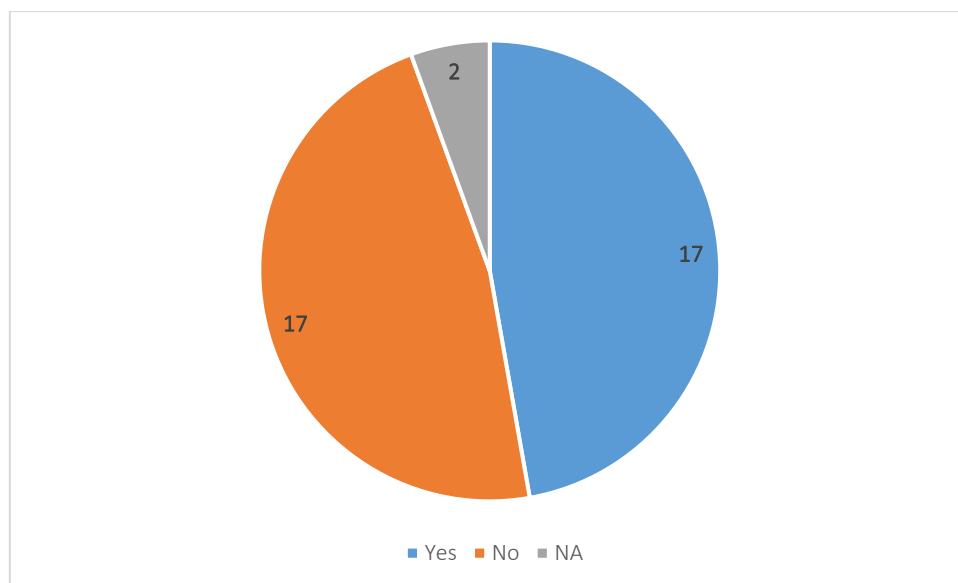


Figure 115: Do you believe that datasets of sufficient quality exist for developing *in silico* methodologies?

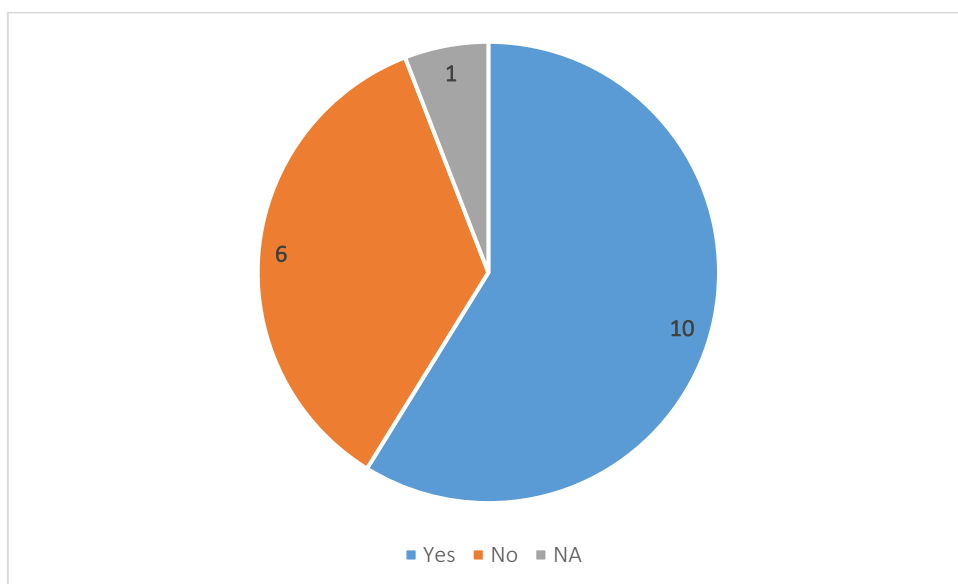


Figure 116: Is it easy to access such datasets (e.g., publicly available)? This graph presents the answers of the experts that believe that datasets of sufficient quality exist for developing *in silico* methodologies.



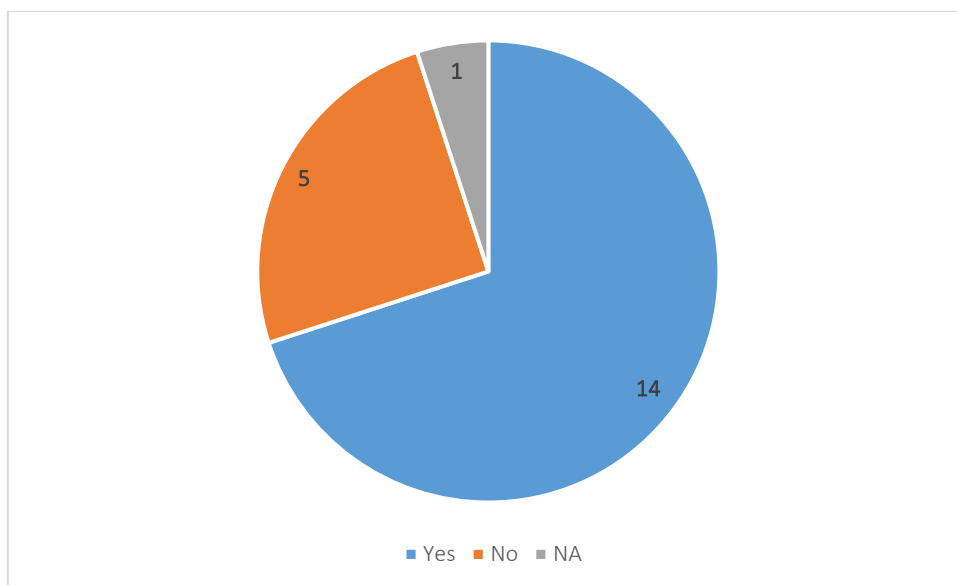


Figure 117: †Do you include data quality control as a standard procedure prior to any modelling activities?

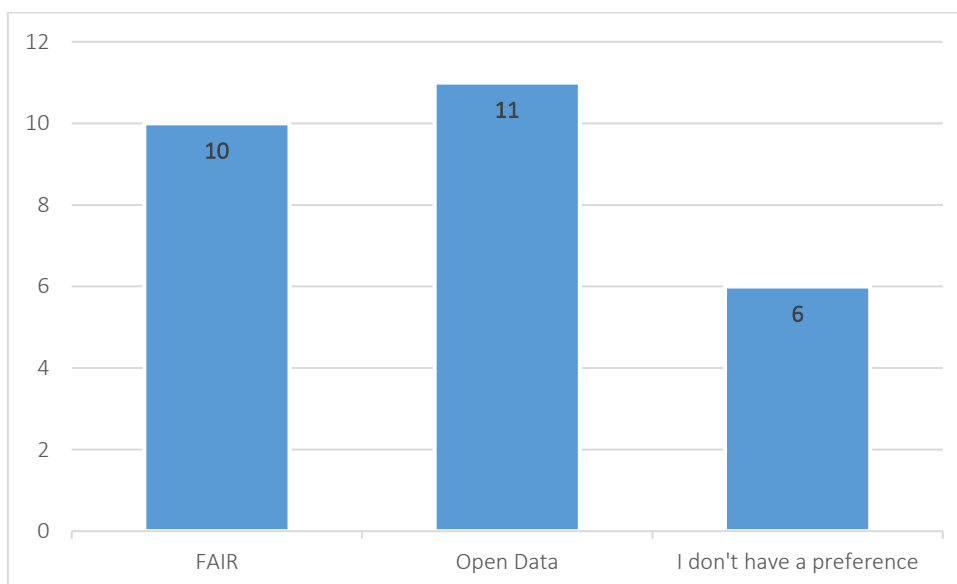


Figure 118: †Do you prefer to use data from sources compliant with the FAIR and/or Open Data guidelines over other sources of data?

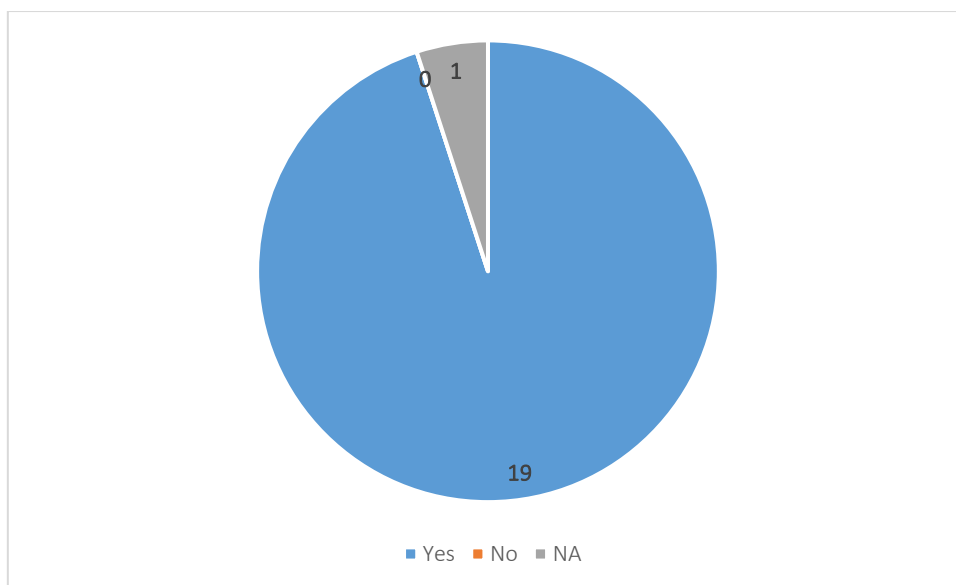


Figure 119: †In your opinion, should data collation from different sources be performed in case of data scarcity?

#### Section 4 – Dissemination

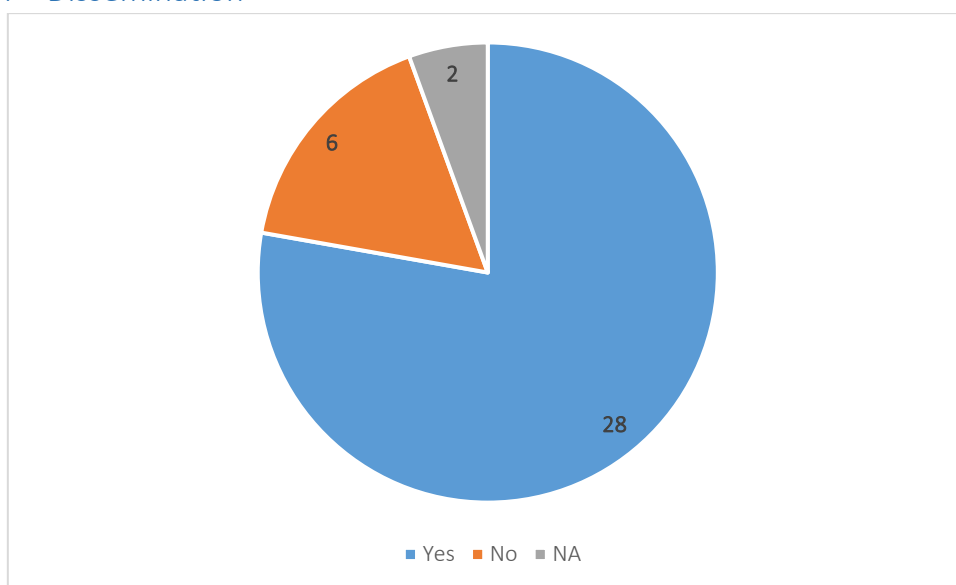


Figure 120: In your opinion is the dissemination of the developed *in silico* methodologies as user-friendly workflows necessary to support the regulatory applications and the SbD of novel nanomaterials?

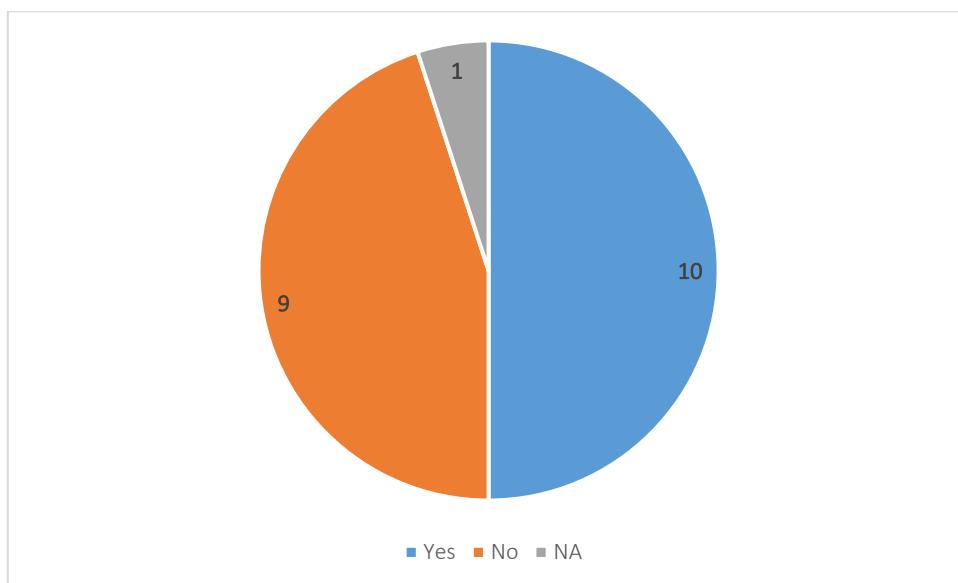


Figure 121: †Is there a licensing system for using your tools?

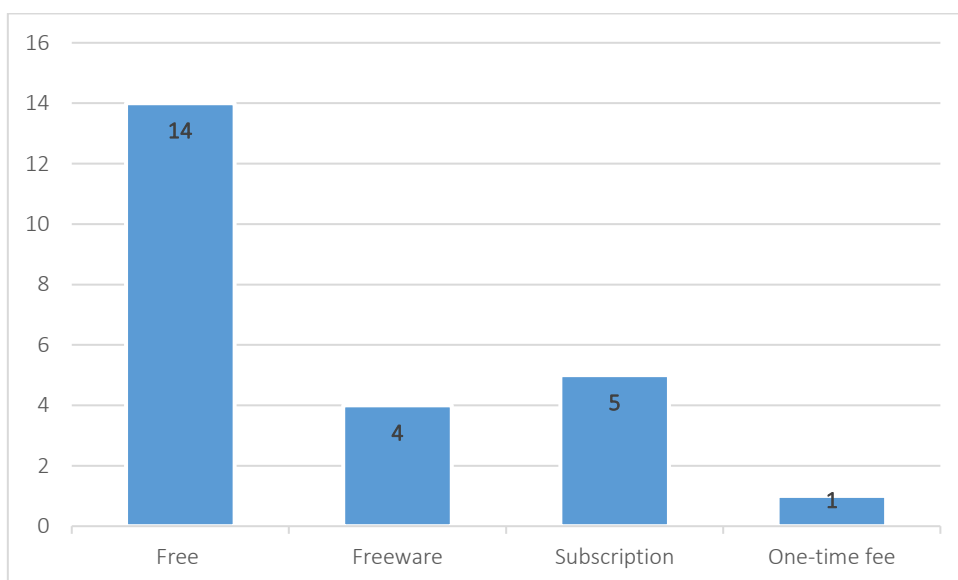


Figure 122: †Are your models freely available or is there a fee for using?

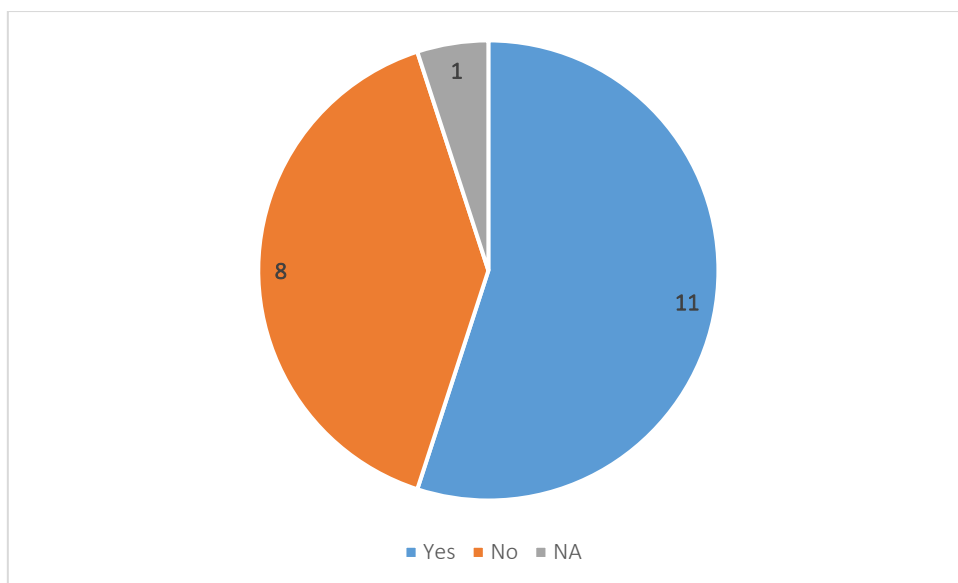


Figure 123: †Are you providing the source code of the tools you are developing?

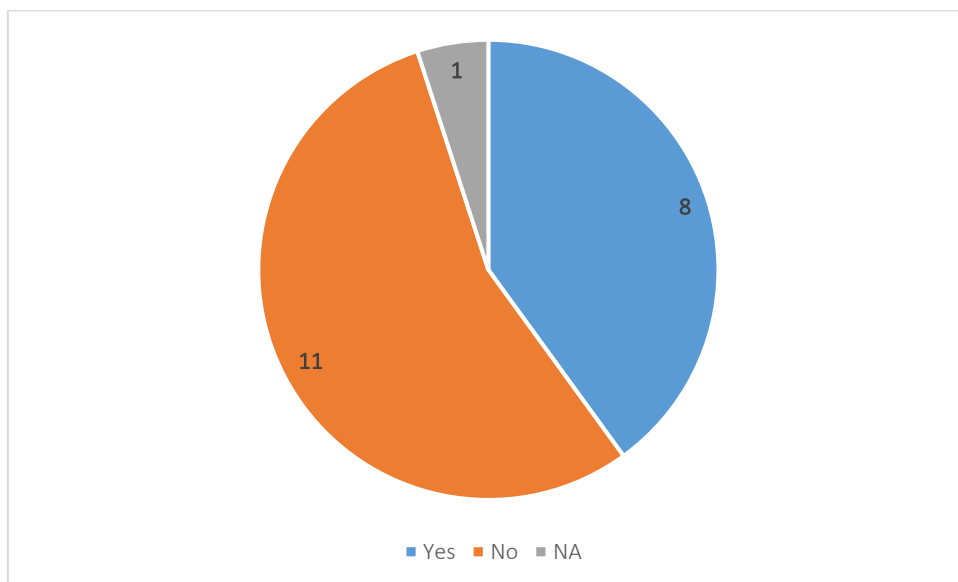


Figure 124: †Do you provide APIs for remote accessing of the *in silico* tools?

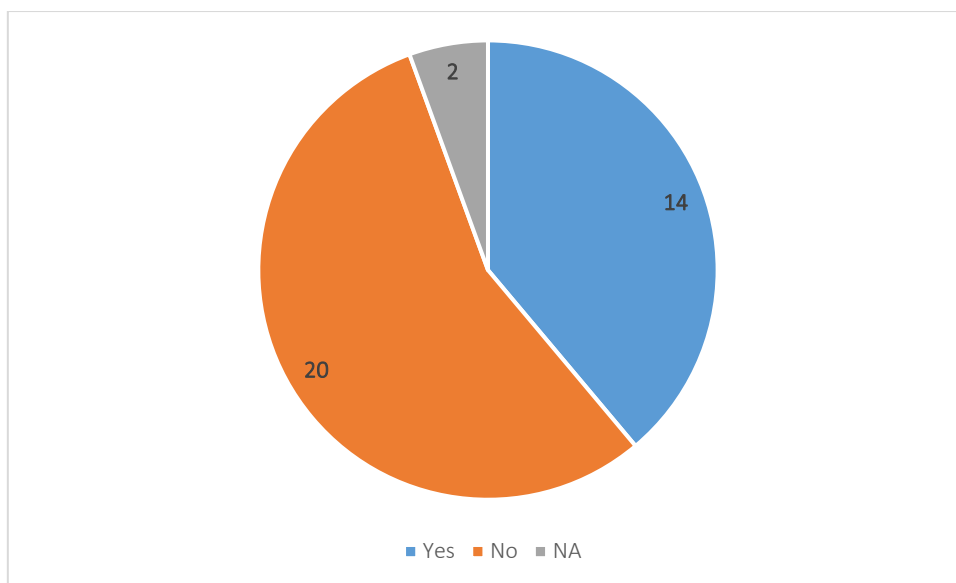


Figure 125: Are you aware of any *in silico* methodologies integrated in industrial, regulatory or research level for the SbD, grouping, and read-across of nanomaterials?

## Section 5 – Gaps and future steps

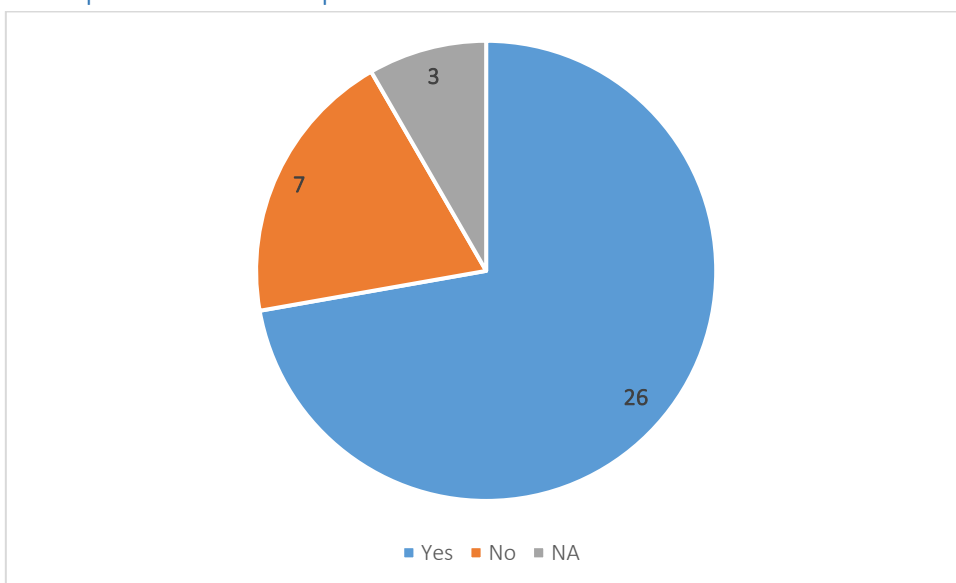


Figure 126: Is it possible to integrate the results of the *in silico* methodologies in the SbD of novel nanomaterials or regulatory applications?

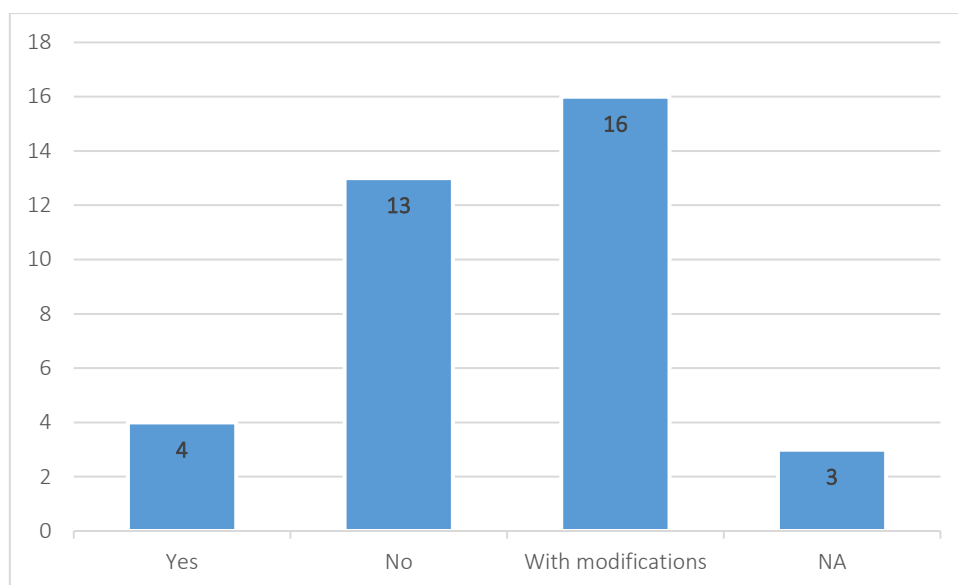



Figure 127: Can *in silico* methodologies used for bulk chemicals be used in the case of nanomaterials?

## Appendix V – QMRF report for the EnaloskNN model (Case study 1)

	<p><b>QMRF identifier (JRC Inventory):</b> To be entered by JRC.</p> <p><b>QMRF Title:</b> Read-across model for the assessment of interactions between carbon nanoparticles and a SARS-CoV-2 RNA fragment.</p> <p><b>Printing Date:</b> 26-Jan-2023</p>
---	--

### 1. QSAR identifier

#### 1.1. QSAR identifier (title):

Read-across model for the assessment of interactions between carbon nanoparticles and a SARS-CoV-2 RNA fragment.

#### 1.2. Other related models:

#### 1.3. Software coding the model: KNIME

### 2. General information

#### 2.1. Date of QMRF:

13/01/2023

#### 2.2. QMRF author(s) and contact details:

#### 2.3. Date of QMRF update(s):

#### 2.4. QMRF update(s):

#### 2.5. Model developer(s) and contact details:

#### 2.6. Date of model development and/or publication:

Model Developed on 20/11/2022.

#### 2.7. Reference(s) to main scientific papers and/or software package:

#### 2.8. Availability of information about the model:

#### 2.9. Availability of another QMRF for exactly the same model:

### 3. Defining the endpoint - OECD Principle 1

#### 3.1. Species:

Different types of carbon-based nanoparticles

#### 3.2. Endpoint:

Total Potential Energy (Interaction energies,  $E_{int}$ )

#### 3.3. Comment on endpoint:

The endpoint has numerical values originally derived via Molecular Dynamics Simulations. The obtained energies had negative values, a reflection of favourable interactions, indicating that CNPs form stable complexes with the RNA fragment.

#### 3.4. Endpoint units:

kJ/mol

#### 3.5. Dependent variable:

Interaction energy ( $E_{int}$ )

#### 3.6. Experimental protocol:

### 3.7. Endpoint data quality and variability:

#### 4. Defining the algorithm - OECD Principle 2

##### 4.1. Type of model:

Machine learning regression model, *k*-Nearest Neighbours used within the read-across framework.

##### 4.2. Explicit algorithm:

*k*-Nearest Neighbours algorithm

*k*=3, inverted distance as weighting factor for the *k* nearest points

##### 4.3. Descriptors in the model:

[1] Molecular Weight g/mol

[2] Overall Surface Area (nm<sup>2</sup>)

[3] Specific Surface Area (nm<sup>3</sup>)

[4] Sum of Degrees dimensionless Topological, SDeg of every carbon atom

[5] Structure Categorical, C-type family each NP belongs to

##### 4.4. Descriptor selection:

Feature selection according to the correlation between all pairs of variables, Spearman's rank correlation coefficient was used to assess the strength and direction of the relationships. Threshold value of 90% correlation filtered out the "Volume" descriptor of the original dataset from further analysis.

##### 4.5. Algorithm and descriptor generation:

Descriptors in the original publication were obtained via molecular dynamics simulations, and from the Multiwfn 3.8 and the Chem3D Ultra (Ver. 19.0) software.

##### 4.6. Software name and version for descriptor generation:

##### 4.7. Chemicals/Descriptors ratio:

17:5

#### 5. Defining the applicability domain - OECD Principle 3

##### 5.1. Description of the applicability domain of the model:

The applicability domain is defined by fixed boundaries (APD threshold), calculated by considering Euclidean distances between all CNPs in the training set.

##### 5.2. Method used to assess the applicability domain:

The distance of a test NP to its nearest neighbour in the training set is compared to the predefined applicability domain threshold (APD), calculated as  $APD = 'd' + 'stdev' * z$ . First, the average of Euclidean distances between all pairs of training compounds is calculated and then the set of distances that were lower than the average is formulated. 'd' and 'stdev' are finally calculated as the average and standard deviation of all distances included in this (remaining) set. ( $z=0.5$ , empirical parameter). If the similarity is beyond this threshold, the prediction is considered unreliable.

##### 5.3. Software name and version for applicability domain assessment:

Domain-APD, Enalos+ node

Enalos+ extension in KNIME analytics platform

##### 5.4. Limits of applicability:



APD threshold=1.963

#### **6.Internal validation - OECD Principle 4**

##### **6.1.Availability of the training set:**

Yes

##### **6.2.Available information for the training set:**

CAS RN: No

Chemical Name: Yes

Smiles: No

Formula: Yes

INChI: No

MOL file: No

##### **6.3. Data for each descriptor variable for the training set: All**

##### **6.4. Data for the dependent variable for the training set: All**

##### **6.5. Other information about the training set:**

CNPs belong to three C-type families (fullerenes, carbon nanotubes, and graphene sheets), Kennard-Stone algorithm used for partitioning, 12 out of 17 CNPs were included in the training set which was used for model development.

##### **6.6. Pre-processing of data before modelling:**

Gaussian normalisation of descriptors (z-score)

##### **6.7. Statistics for goodness-of-fit:**

Squared correlation coefficient ( $r^2$ ), Mean absolute error (MAE), Root mean square error (RMSE)

##### **6.9. Robustness - Statistics obtained by leave-many-out cross-validation:**

##### **6.10.Robustness - Statistics obtained by Y-scrambling:**

##### **6.11. Robustness - Statistics obtained by bootstrap:**

##### **6.12. Robustness - Statistics obtained by other methods:**

Golbraikh and Tropsha's test, the model satisfies all their conditions.

#### **7.External validation - OECD Principle 4**

##### **7.1. Availability of the external validation set:**

Yes

##### **7.2. Available information for the external validation set:**

CAS RN: No

Chemical Name: Yes

Smiles: No

Formula: Yes

INChI: No

MOL file: No

**7.3. Data for each descriptor variable for the external validation set:**

All

**7.4. Data for the dependent variable for the external validation set:**

All

**7.5. Other information about the external validation set:**

5 out of 17 CNPs were included in the test set. The test set was not involved in model development, it was used solely for validating purposes.

**7.6. Experimental design of test set:**

The initial dataset was split into 70:30 ratio for training and test subsets respectively using the Kennard-Stone algorithm.

**7.7. Predictivity - Statistics obtained by external validation:** Squared correlation coefficient ( $r^2$ )= 95%, Mean absolute error (MAE)=15.4, Root mean square error (RMSE)=20.4

**7.8. Predictivity - Assessment of the external validation set:**

The external validation set is 30% of the initial dataset and all predictions fall within the DoA.

**7.9. Comments on the external validation of the model:**

Gaussian normalisation on descriptors was applied based on the training set.

**8. Providing a mechanistic interpretation - OECD Principle 5**

**8.1. Mechanistic basis of the model:**

Molecular weight characterises the molecule's size, CNP Structure was considered since the interaction affinity with the cov-RNA fragment increases in the order of fullerenes < graphene sheets < C-nanotubes. SDEg is a topological descriptor represents the total number of vertex degrees of non-hydrogen atoms connected to every atom of the CNP. Also, CNPs have high surface areas which allows good adsorption and desorption behaviour; therefore, surface descriptors (OSA and SSA) were included (The larger the surface area of the nanomaterial, the greater the number of adsorption sites available for a molecule to bind to). In the original publication, each CNP's "Volume" was used for obtaining their surface area values, which explains the descriptor's high correlation with other variables and its exclusion from the dataset.

**8.2. A priori or a posteriori mechanistic interpretation:**

Posteriori-After modelling, by interpretation of the final set of descriptors

**8.3. Other information about the mechanistic interpretation:**

**9. Miscellaneous information**

**9.1. Comments:**

The original publication was "Probing nano-QSAR to assess the interactions between carbon nanoparticles and a SARS-CoV-2 RNA fragment", by Zhang *et al.* (2021). In this publication three separate models were developed for the prediction of  $E_{int}$ , our model had a slightly better performance compared to them.

**9.2. Bibliography:**

[1] Probing nano-QSAR to assess the interactions between carbon nanoparticles and a SARS-CoV-2 RNA fragment <https://doi.org/10.1016/j.ecoenv.2021.112357>.

### 9.3. Supporting information:

Training set(s) Test set(s) Supporting information

## 10. Summary (JRC QSAR Model Database)

### 10.1. QMRF number:

To be entered by JRC

### 10.2. Publication date:

To be entered by JRC

### 10.3. Keywords:

To be entered by JRC

### 10.4. Comments:

To be entered by JRC

## Appendix VI – Quantitative measures of goodness-of-fit and predictivity for regression models

Most of the statistical metrics presented here are also included in Table 5.

### Correlation of determination ( $R^2$ )

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad [1]$$

Where  $N$ , is the number of samples,  $y_i$  and  $\hat{y}_i$ , are the actual and predicted endpoint values of the  $i^{th}$  sample respectively, and  $\bar{y}$  and  $\bar{\hat{y}}$ , are the average endpoint values of the experimental and predicted values respectively.

### External explained variance ( $Q^2_{\text{ext}}$ or $Q^2_{F1}$ )

$$Q^2_{\text{ext}} \text{ or } Q^2_{F1} = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_{\text{train}})^2} \quad [6]$$

$$Q^2_{F2} = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_{\text{test}})^2} \quad [7]$$

Where  $N$ , is the number of test samples,  $y_i$  and  $\hat{y}_i$ , are the actual and predicted endpoint values of the  $i^{th}$  test sample respectively.  $\bar{y}_{\text{train}}$  and  $\bar{y}_{\text{test}}$  are the average actual activities of training set and test set samples respectively.

### Mean absolute error (MAE)

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad [3]$$

Where  $N$ , is the number of samples,  $y_i$  and  $\hat{y}_i$ , are the actual and predicted endpoint values of the  $i^{th}$  sample respectively.

### Mean square error (MSE)

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad [4]$$

Where  $N$ , is the number of samples,  $y_i$  and  $\hat{y}_i$ , are the actual and predicted endpoint values of the  $i^{th}$  sample respectively.

### Root mean square error (RMSE)

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad [5]$$

Where  $N$ , is the number of samples,  $y_i$  and  $\hat{y}_i$ , are the actual and predicted endpoint values of the  $i^{th}$  sample respectively.

### Squared Pearson correlation coefficient ( $r^2$ )

$$r^2 = \left( \frac{\sum_{i=1}^N (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^N (y_i - \bar{y})^2 \sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}})^2}} \right)^2 \quad [10]$$

Where  $N$ , is the number of samples,  $y_i$  and  $\hat{y}_i$ , are the actual and predicted endpoint values of the  $i^{th}$  sample respectively, and  $\bar{y}$  and  $\bar{\hat{y}}$ , are the average endpoint values of the experimental and predicted values respectively.

### Golbraikh and Tropsha's test

A series of statistical criteria that an acceptable regression QSAR model should satisfy according to Golbraikh and Tropsha<sup>252</sup>.

Statistic	Rule
$r^2$	$> 0.6$
$Q_{\text{ext}}^2$	$> 0.5$
$\frac{r^2 - R_0^2}{r^2}$ or $\frac{r^2 - R_0'^2}{r^2}$	$< 0.1$
$k$ or $k'$	$\in [0.85, 1.15]$
$ R_0^2 - R_0'^2 $	$< 0.3$

Where:

$$k = \frac{\sum_{i=1}^N y_i \hat{y}_i}{\sum_{i=1}^N \hat{y}_i^2} \quad [23]$$

$$k' = \frac{\sum_{i=1}^N y_i \hat{y}_i}{\sum_{i=1}^N y_i^2} \quad [24]$$

$$R_0^2 = 1 - \frac{\sum_{i=1}^N (\hat{y}_i - \hat{y}_i^{r_0})^2}{\sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}})^2}, \text{ where } \hat{y}_i^{r_0} = k' y \quad [25]$$

$$R_0'^2 = 1 - \frac{\sum_{i=1}^N (y_i - y_i^{r_0})^2}{\sum_{i=1}^N (y_i - \bar{y})^2}, \text{ where } y_i^{r_0} = k \hat{y} \quad [26]$$

Where  $N$ , is the number of samples,  $y_i$  and  $\hat{y}_i$ , are the actual and predicted endpoint values of the  $i^{th}$  sample respectively, and  $\bar{y}$  and  $\bar{\hat{y}}$ , are the average endpoint values of the experimental and predicted values respectively.