The use of computational toxicology in hazard assessment of engineered nanomaterials

Chen, G.; Chen G.

The use of computational toxicology in hazard assessment of engineered nanomaterials

陈广超

Guangchao Chen

The use of computational toxicology in hazard assessment of engineered nanomaterials.

Ph.D. Thesis Leiden University, The Netherlands

ISBN: 978-94-6182-823-1

Cover design: Guangchao Chen

Printed by: Off Page, www.offpage.nl

The use of computational toxicology in hazard assessment of engineered nanomaterials

Proefschrift

ter verkrijging van de graad van

Doctor aan de Universiteit Leiden,

op gezag van de Rector Magnificus Prof. mr. C.J.J.M. Stolker

volgens besluit van het College van Promoties

te verdedigen op 19 september 2017

klokke 13:45 uur.

door

Guangchao Chen

Geboren te Qiqihar, China

In 1987

Promotiecommissie:

Promotor:	Prof. dr. W.J.G.M. Peijnenburg
Co-promotor:	Dr. M.G. Vijver
Overige leden:	Prof. dr. ir. D. van de Meent (RIVM)
	Dr. K. Jagiello (University of Gdansk)
	Dr. I. Tetko (Helmholtz Zentrum München)
	Prof. dr. Arnold Tukker (Universiteit Leiden)
	Prof. dr. ir. P.M. van Bodegom (Universiteit Leiden)

Table of Contents

CHAPTER 1

GENERAL INTRODUCTION

1.1 Nanotechnology and nanomaterials

1.1.1 Background

Nanotechnology has become a trending topic in the 21st century. It basically deals with controlling the structure of matter at the nanoscale (1-100 nm) with respect to one or more external dimensions in order to produce new materials, i.e. nanomaterials (Maynard et al., 2006; European Commission, 2011). The prefix 'nano' originates from the Greek word for "dwarf" (Boholm, 2016). As a prefixing unit of time, length, mass etc., nano signifies "a billionth" (e.g. of a meter or a gram). Therefore one nanometer (nm) is equal to one-billionth of a meter, i.e. 10^{-9} m. To put this scale into perspective, a human hair is about 80,000 nm wide; a DNA molecule is around 2.5 nm wide; and a red blood cell is estimated to be approximately 7,000 nm wide (Sahoo et al., 2007; Thakkar et al., 2010).

Nanomaterials are tailored to the needs of inimitable characteristics (e.g. electromagnetic, catalytic, optical, and thermal properties) which are often not observed in their bulk counterparts (Kleandrova et al., 2014; Puzyn et al., 2009). As such, nanomaterials have been designed and engineered for a broad spectrum of applications. An online database called the "Nanotechnology Consumer Products Inventory" has listed eight general categories (including 37 sub-categories) of nano-enabled products, namely appliances, automotive, cross cutting, electronics and computers, food and beverage, goods for children, health and fitness, and home and garden (Project on Emerging Nanotechnologies, 2013; Vance et al., 2015). By 5 December, 2016, this inventory contained in total 1827 consumer products on the market from 715 companies in 33 countries, which were manufacturer-identified as incorporating engineered nanomaterials (ENMs). Meanwhile by the same date, another data and analysis repository named the "Nanotechnology Products Database" provided a total number of 6396 nanotechnology products introduced by 910 companies in 49 countries (Nanotechnology Products Database, 2016). The controllable production and widespread commercial applications of ENMs have shown the immense promise of nanomaterials to benefit the world economy and quality of life. As reported, the direct employment in the EU involving nanotechnology is estimated to be up to 400,000 jobs according to the European Commission (Lynch, 2016).

1.1.2 Rapid development of nanotechnology

Reportedly, every week about 3 to 4 new ENM-incorporated products are likely to enter the market (Kar et al., 2014). The worldwide production capacity of ENMs is estimated to increase from only 2000 tons per year in 2004 to 50,000 tons per year by 2020 (Heggelund et al., 2014). As predicted, the global production rates for ENMs involved in structural

applications (e.g. catalysts, films & coatings, composites) will see an increase from an order of 10^3 tons per year in 2010 to an order of 10^4-10^5 tons per year by 2020 (Dowling et al., 2004; Borm et al., 2006). During the same period, the global production of ENMs applied in information and communication technologies is expected to increase from 10^2 to $>10^3$ tons per year; global production volume of ENMs used for environmental applications (such as nanofiltration, membranes) is expected to rise from 10^2 to 10^3-10^4 tons per year; worldwide production of ENMs for skincare products (e.g. TiO_2, ZnO ENMs) is estimated to see a steady growth of 10^3 tons per year (Dowling et al., 2004). By 2020, the global market of nanotechnology is likely to continue to grow at double-digit rates (around 17% annually) for the coming decade and reach a global value of $75 billion in 2020 (Mulvaney and Weiss, 2016).

1.2 Safety concerns of nanomaterials

1.2.1 Release of ENMs into the environment

The rapid development of nanotechnology and extensive use of ENMs for industrial and commercial applications have caused safety concerns (Nel et al., 2006; Valsami-Jones and Lynch, 2015). The ongoing production of ENMs of all types certainly increases the likelihood of the release of ENMs into the environment. A proposed life-cycle of ENMs showed that the release of ENMs into the environment can be traced back to the stages of ENM production, incorporation of ENMs into products, and consumption, recycling, and disposal of the ENM-containing products (Gottschalk et al., 2010). As estimated, every year about 189,200 tons of ENMs are released into landfills; 69,200 tons of ENMs into water bodies; 51,600 tons of ENMs into soil; and 8,100 tons of ENMs into the air (Keller and Lazareva, 2014). In Europe, the concentration of Ag ENMs in the air was estimated to be around 0.008 ng/m^3 in 2008 (volume of air in EU was estimated to be 4.33×10^{15} m^3); in surface water (estimated volume in EU 3.89×10^{14} liter) the predicted concentrations are 0.764 ng/L for Ag ENMs, 0.010 μg/L for ZnO ENMs, and 0.015 μg/L for TiO_2 ENMs (Gottschalk et al., 2009; Sun et al., 2014). Since 2008, the annual increases of ENM concentrations in soil (0.05 m depth as for natural soil, 0.2 m depth as for agricultural soil, estimated total volume in EU 7.59×10^{14} kg) are predicted to be 0.0227, 0.093, and 1.28 μg/kg for Ag, ZnO, and TiO_2 ENMs, respectively (Gottschalk et al., 2009; Sun et al., 2014).

1.2.2 Exposure to ENMs

Undoubtedly, the ongoing release of ENMs into the environment inevitably results in a higher exposure of humans and ecosystems to ENMs. As illustrated in Figure 1.1, the exposure of humans to ENMs may occur via a number of exposure routes:

(i) Dermal contact, for instance by applying personal care products incorporating ENMs such as TiO_2 and ZnO ENMs (Keller et al., 2014);
(ii) Inhalation. In an environment where ENMs are released into the air, the manufactured nanomaterials can be inhaled directly and thus get to deposit in the lung (Methner et al., 2010);
(iii) Ingestion. This is due to the ENMs added to food items, or unwarranted ENMs that leach off of package materials into food (Magnuson et al., 2011; McCracken et al., 2016), or the ENM-polluted water (Wang et al., 2008).

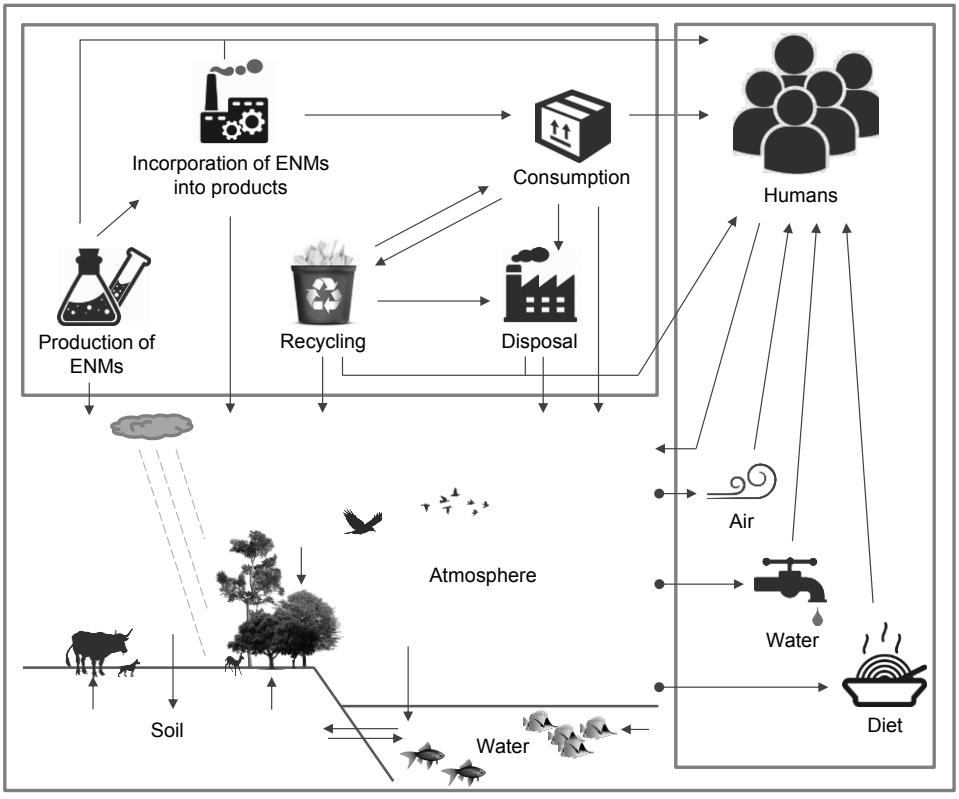


Figure 1.1. Possible exposure routes of humans and the environment to ENMs (adapted from Dowling et al., 2004; Gottschalk et al., 2010).

Meanwhile, therapeutic and medical applications of ENMs can also result in direct uptake of those materials into the human body, even though this option is still seen as underdeveloped (Dowling et al., 2004). Theoretically, the proposed exposure routes of

ENMs also apply for environmental organisms just as for humans. However, given the diversity of organisms that live in the outside world, exposure of ENMs to environmental species seems to be much more complicated, and additional exposure and uptake routes do certainly exist. For instance, the gill is concluded as being the principal site of the uptake of Cu and Ag ENMs for fishes and other gill-keeping species (Kwok et al., 2012; Griffitt et al., 2007). Plants could interact with ENMs adsorbed on soil and sediments via roots (Oberdörster et al., 2005). ENMs deposited and aggregated on the leaves or other aerial parts of plants are able to penetrate through stomatal pathways (Eichert et al., 2008; Miralles et al., 2012). ENMs have shown to be taken up through bacterial cell membranes (Klaine et al., 2008; Kumar et al., 2011).

1.2.3 Potential toxicity of ENMs and possible mechanisms

The uptake of nanomaterials may lead to adverse effects. Previously, consensus was drawn across a majority of studies regarding the occurrence of damage triggered by ENMs at the cellular level (Bondarenko et al., 2013). A comprehensive study of Shaw et al. (2008) evidenced the effects of ENMs on the cellular viability and physiology of different mammal cell lines. Gajewicz et al. (2015) also reported the impacts of metal oxide ENMs on cell viability (human keratinocyte cells) which confirmed the observations of other independent reports (Zhang et al., 2012; Liu et al., 2011; Zhou et al., 2008). Experimental assays of nanotoxicity have also been generally performed on various trophic levels of organisms (Juganson et al., 2015; Donaldson et al., 2001; Oberdorster, 2000), such as algae (e.g. *Pseudokirchneriella subcapitata*), bacteria (e.g. *Escherichia coli*), crustaceans (e.g. *Daphnia magna*, *Daphnia pulex*), fish (e.g. *Danio rerio*, *Oryzias latipes*), nematodes (e.g. *Caenorhabditis elegans*), plants (e.g. *Lemna minor*), protozoa (e.g. *Tetrahymena thermophila*), yeast (e.g. *Saccharomyces cerevisiae*), and mammals (e.g. *Rattus*).

The introduction of ENMs to different species may lead to the occurrence of (not limited to) mortality, immobilization, malformation, inflammatory response, and the inhibition of cell viability, growth, luminescence, reproduction, feeding, and fertilization, etc. which varies from case to case (Juganson et al., 2015). For instance, exposure of algae (e.g. *Chlamydomonas reinhardtii*, *Pseudokirchneriella subcapitata*, and *Scenedesmus obliquus*) to ENMs may result in the inhibition of growth and the loss of cell viability (Navarro et al., 2008; Angel et al., 2013; Dalai et al., 2013). Bacteria that were exposed to ENMs appeared to exhibit mortality (e.g. *Escherichia coli*), luminescence inhibition (e.g. *Escherichia coli*, *Vibrio fischeri*, *Pseudomonas putida*), and growth inhibition (e.g. *Escherichia coli*) (Samberg et al., 2011; Ivask et al., 2010; Hu et al., 2009; Dams et al., 2011; Heinlaan et al., 2008). Exposure of crustaceans, such as *Daphnia magna*, *Ceriodaphnia dubia*, and *Daphnia pulex* to ENMs was found to cause mortality, immobilization, and inhibition of growth, feeding, and reproduction (Gao et al., 2009; Li et

al., 2011; Jo et al., 2012; Lopes et al., 2014; Griffitt et al., 2008). ENMs tested on fish are likely to induce mortality, growth inhibition, delay of hatching, and developmental malformation (Wang et al., 2012; Massarsky et al., 2013; Zhu et al., 2012; Hall et al., 2009). The introduction of ENMs to nematodes, protozoa, and yeast may cause mortality, growth inhibition, inhibition of cell viability, reproduction inhibition, and immobilization (Tyne et al., 2013; Yang et al., 2012; Ma et al., 2009; Kvitek et al., 2009; Shi et al., 2012; Mortimer et al., 2010; Galindo et al., 2013; Kasemets et al., 2013). For rats, *in vivo* experiments have also evidenced the harmful effects of ENMs such as hepatotoxicity and nephrotoxicity after oral gavage (Lei et al., 2008).

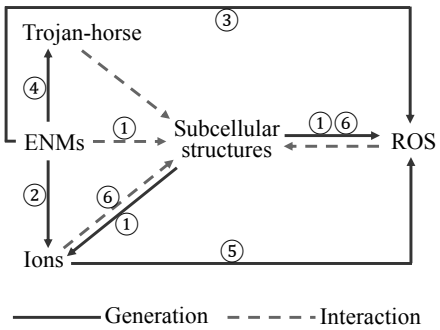


Figure 1.2. Schematic illustration of possible mechanisms of ENM toxicity. 1) The direct contact of ENMs with subcellular structures which could promote the leaching of ions and reactive oxygen species (ROS); 2) ENMs releasing ions; 3) ENMs contact-mediated ROS generation; 4) The phenomenon of Trojan-horse mechanism; 5) Released ions enhancing the formation of ROS; 6) Ion-dependent interactions which may result in cellular damage or trigger ROS formation (Drawn by G. Chen).

As hypothesized, ENMs may pose effects via a single or via combinations of a few possible pathways (see Figure 1.2). ENMs can for instance induce the generation of reactive oxygen species (ROS), or induce direct steric hindrance or interferences with important reaction sites (Puzyn et al., 2011). ENMs are also considered to be able to act as vectors for transporting other toxic chemicals into cells, a phenomenon which is described as the Trojan-horse mechanism (Park et al., 2010). Nanotoxicity could as well occur due to the shedding of ions from ENM crystals. This process is generally believed to be one of the important pathways of toxicity for soluble metallic nanoparticles (Xiao et al., 2015). The released ions are able to interact with subcellular structures initiating cellular damages, or stimulate ROS formation, which in turn has been reported to induce oxidative stress

resulting in the disturbance of cellular physiological redox-regulated functions (Nel et al., 2009; Fu et al., 2014). The released ions can also promote the production of ROS (von Moos and Slaveykova, 2014).

1.3 Handling nanosafety

1.3.1 Environment risk assessment and safe-by-design of ENMs

To ensure the nanosafety and optimal benefit from nanotechnology, two strategies stand out in this regard: the first is designing and producing ENMs that are safe and environmentally benign while with desired properties. This strategy aims at minimizing the potential risks of ENMs from the very beginning of the development of an ENM application. It is referred to as the safe-by-design of ENMs and relates to *ex ante* safety assessment. The second strategy is to assess the risks of existing and also newly introduced ENMs, which manages to control relevant risks during the stages of manufacturing, use, and disposal of ENMs with prospective risk assessment. In order to answer the question whether an ENM is environmentally safe or not, a series of key steps including hazard assessment, exposure assessment, and risk characterization are required, and measures will be taken based on the established conclusion (Commission of the European Communities, 1996). As suggested by the European Chemicals Agency (ECHA), the process of risk assessment of ENMs (as of any chemical to be regulated within the EU) begins with the identification and assessment of ENM hazard if the ENMs subjected to registration under the Registration, Evaluation, Authorization and Restriction of Chemicals (REACH) regulation reaches an annual production or import of at least 10 tons (ECHA, 2011). Exposure assessment and risk characterization are also required when an ENM fulfils the criteria for any of the listed physical, health, or environmental hazard categories in the released guideline of ECHA (ECHA, 2011). The hazard assessment of ENMs meanwhile also provides important feedbacks to safe-by-design approaches for ENMs with regard to key characteristics of ENMs governing relevant toxicity pathways, upon which modifications towards designing safer materials could be determined (Sealy, 2011).

As described in Figure 1.3, the hazard assessment for human health and the environment as recommended by ECHA comprises of evaluation of information, classification and labelling, and identification of predicted no effect concentrations (PNECs) or derived no effect levels (DNELs). The gathering and evaluation of relevant physicochemical, (eco)toxicological information of ENMs is certainly fundamental to support the assessment of ENM hazard as the very first step. This is outlined by REACH as part of the registration of chemicals. It

involves the retrieval and sharing of existing data, consideration of needed information, identification of information gaps, and the generation of new data or preparation of a proposal for a tailored testing strategy (ECHA, 2011). Based on the first step, ENMs will be determined as whether or not meeting the criteria for any of the hazard classes or categories proposed by ECHA, i.e. the step of classification and labeling. Once an ENM is categorized in at least one of the listed classes, derivation of the hazard threshold levels of ENM for human health and the environment, e.g. PNECs and DNELs, is required in light of a qualitative risk characterization for relevant ENMs.

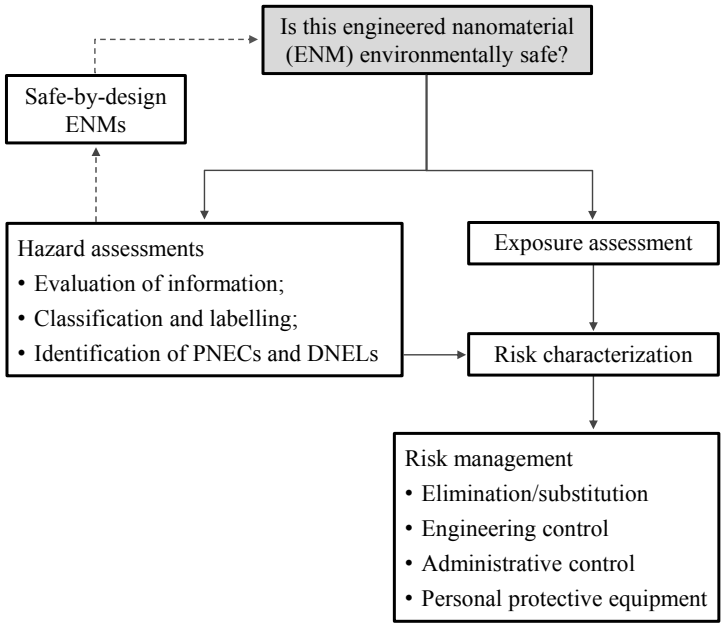


Figure 1.3. Schematic explanation of the safe handling of engineered nanomaterials. PNECs - predicted no effect concentrations; DNELs - derived no effect levels. Figure adapted from the ECHA guidance (ECHA, 2011).

Therefore, as for the safe handling of nanomaterials including the approaches of safe-by-design and risk assessment of ENMs, gathering and evaluation of hazard information of ENMs is essential. In a survey provided by the NanoSafety Cluster Database working group, a total number of 38 online ENM databases developed under various projects were listed. These were provided with the names of databases and website addresses (Mustad et al., 2014). Another online inventory named StatNano was established in 2010 for the access of

up-to-date information and statistics in nano-based science, technology and industry (StatNano, 2010). This website also provides a so-called Nanotechnology Products Database (established in January 2016) for the analysis and characterization of nanotechnology-based consumer products (Nanotechnology Products Database, 2016). Another effort to gather relevant data and to address the safety of ENMs is the EU NanoSafety Cluster which aims to maximize the synergies between various projects at the European-level (EU NanoSafety Cluster, 2017). This cluster comprises of nine working groups addressing different aspects involved in nanotechnology and nanosafety, namely the working groups of materials, hazard, exposure, database, risk, modeling, dissemination, systems biology, and safe-by-design and industrial innovation. Undoubtedly, those databases and platforms are of significant importance as the first step in gathering, evaluation, and processing of information regarding the hazard of ENMs. As a follow up, it is crucial to develop comprehensive databases containing reliable and sufficient information on ENM characterization, experimental conditions, and toxicity of ENMs for the need of ENM hazard assessment.

1.3.2 Handling nanosafety with the aid of computational toxicology

By far, a large amount of ENMs have been carefully tested on various species and cell lines (Bondarenko et al., 2013; Juganson et al., 2015). However, given the substantial number of existing, non-tested ENMs and the enormous growth of nanotechnology, testing every single type of ENM to support the comprehensive evaluation of ENM safety is expensive, time-consuming, and thus virtually impossible. Testing of all hitherto non-tested ENMs and of all newly developed ENMs also conflicts with the 3R's principle (refine, reduce, and replace) of animal use in toxicity testing (Russell and Burch, 1959). Thus researchers have been seeking and developing alternatives of testing assays for assisting the risk management of ENMs. One of the very helpful tools in this task as an alternative of testing is computational toxicology. It is defined as a discipline that integrates information from various sources in order to develop computer-based models for the better interpretation and prediction of chemical effects (Reisfeld and Mayeno, 2012). A few typical tools in this field are for example (quantitative) structure–activity relationship ((Q)SAR), structural alerts, read-across extrapolations, dose–response and time–response models which aim to contribute to the prediction and classification of chemical toxicity.

The (Q)SAR method enables the correlation of chemical characteristics with experimental toxicity data and thus enables to encode existing knowledge into predictive models. To build a (Q)SAR model, the measured or calculated descriptors characterizing key structures of chemicals and the toxicity endpoints reflecting the chemical biological effects are required. The role of (Q)SARs in predictive toxicology is:

(i) To provide efficient and inexpensive screening tools for the evaluation of chemical hazards;

(ii) To assist the categorization and labeling of chemicals based their hazard effects;

(iii) To help interpreting the underlying toxicity mechanisms of substances (Peijnenburg, 2009).

1.3.3 Hazard prediction by (Q)SARs for ENMs

(Q)SARs have already been successfully used as very helpful tools for conventional chemicals in relating structural characteristics to chemical properties and biological effects in order to fill data gaps (Chen et al., 2014; Singh et al., 2014; Modarresi et al., 2007). According to REACH, data derived from (Q)SARs may support the waiving of laboratory testing or serve as a trigger for proposing further testing; when certain required conditions are met, (Q)SAR results could be used instead of testing data for the registration under REACH (ECHA, 2008a). Based on the OECD principles for (Q)SAR validation, a (Q)SAR model suited for regulatory purposes is suggested to contain at least the following information:

(i) A well-defined endpoint;

(ii) An explicit algorithm;

(iii) A well-defined applicability domain;

(iv) Suitable measures of goodness-of-fit, robustness and predictivity;

(v) If possible an interpretation of relevant mechanisms (OECD, 2007).

The limited data availability on ENM hazards necessitates the need of extending conventional (Q)SAR approaches to nanotoxicology, i.e. nano-(Q)SARs. For the hazard assessment of ENMs, nano-(Q)SARs could be potentially used to generate non-testing data during the gathering of information in the first step, or to assist the second step of classification and labeling of ENMs by directly categorizing ENMs into different hazard classes. The descriptors in nano-(Q)SARs may also be helpful for understanding related mechanisms and identifying key factors affecting ENM toxicity, which as well provides guidance to the modification of ENM characteristics for the safe-by-design of ENMs. To date, attempts have already been made to correlate the characteristics of ENMs to their

biological responses (Sizochenko and Leszczynski, 2016; Raies and Bajic, 2016; Tantra et al., 2015). Those studies showed the tantalizing possibility that the (Q)SAR method may indeed be feasible and useful in predicting the biological activity profiles of novel ENMs. However, it meanwhile also revealed that nano-(Q)SAR is now still in its infancy and further challenges in this field need to be overcome. One issue standing out on this background relates to the comprehensive representation of ENM structures. As known, ENMs often exist as populations of materials varying in structural characteristics, e.g. composites, sizes, shapes, functional groups. The structural ambiguousness of ENMs makes it difficult for experimentalists to provide precise information on ENM characterization which consequently hinders the calculation of representative descriptors for ENMs. Another issue of importance in this context concerns the dynamics of ENMs in media. ENMs often strongly interact with constituents in the medium and undergo dramatic changes to their surface properties, and dissolution and aggregation behavior (Winkler, 2016). These changes consequently alter the mobility, bioavailability, and ultimately, toxicity of ENMs. Therefore in some cases the toxicity information of ENMs can be poorly correlated to ENMs' characteristics without considering the dynamics of ENMs in the media. Thus (Q)SARs based on initial structural features of ENMs are now also extended incorporating the experimental descriptors for this consideration (Liu et al., 2011; Zhang et al., 2012).

1.3.4 Hazard prediction models such as SSDs for ENMs

Meanwhile, to derive hazard threshold levels of (soluble) chemicals such as PNEC for ecosystems and their communities, the species sensitivity distribution (SSD) method is commonly used (Posthuma et al., 2002). SSDs are derived by ranking species according to their sensitivity to certain chemicals based on retrieved ecotoxicity data (Posthuma et al., 2002; Garner et al., 2015). An SSD can provide the potentially affected fraction of species under a chemical concentration of interest given the sensitivity distributions. Among others, the 5th percentile (HC5) of the SSD is commonly used to assist in getting protection levels in the ecosystem. The PNEC of a chemical is the maximum acceptable concentration in the environment below which unacceptable chemical effects are unlikely to occur (ECHA, 2008b). It is required for the risk characterization once a chemical is classified into the hazard categories listed in Article 14(4) of REACH (ECHA, 2011). Alternatively, PNEC values can be estimated by the assessment factor method (ECHA, 2011).

1.3.5 Risk characterization for ENMs

PNEC is based on ecotoxicity data and is often coupled with an assessment factor. The obtained values of the hazard thresholds together with the predicted environmental concentrations (PECs) of ENMs are commonly utilized for risk characterization. A ratio of

PEC/PNEC greater than or equal to 1 indicates that potential risks are likely to occur, and further assessment is needed; a ratio of PEC/PNEC less than 1 means that risks are not expected. Even though the risk characterization of PEC/PNEC is originally set up for dissolved chemicals, it is now accepted and widely used for the risk characterization of ENMs as well (Gottschalk et al., 2013; Coll et al., 2016).

1.4 Objectives and outlines of this thesis

As indicated above, assessing and managing the risks of ENMs is of significant importance for the advancement of nanotechnology. Two approaches capable of contributing to this crucial task are principle of safe-by-design of ENM and the risk assessment of ENMs, both of which need to be supported by ENMs' hazard assessment. Computational toxicology as a promising tool has shown its great potential in assisting the evaluation of the hazard of conventional chemicals, from the very beginning of assembling and evaluating data, to classification and labeling, and to generation of hazard threshold values for risk characterization. The use of computational toxicology in supporting the hazard assessment of ENMs is still a field of research that needs further development. This PhD study aims to explore the use of computational toxicology to contribute to the safe handling of metal-based ENMs, by evaluating the availability of existing nanotoxicity data and identifying data gaps, developing nano-(Q)SARs, and deriving hazard threshold values. The objectives of this PhD thesis are:

(i) To evaluate the currently existing literature data on metal-based ENMs for the use of computational toxicology in light of the safety assessment of ENMs;

(ii) To develop nano-(Q)SARs for the prediction and categorization of ENM hazard;

(iii) To derive SSDs and maximum acceptable environmental concentrations of metal-based ENMs as toxicity measures characterizing relevant risks.

1.5 Outline of the thesis

On the basis of the presented research objectives, this thesis contains six chapters.

Chapter 1 Background information is presented about the development of nanotechnology, concerns of nanosafety, safe handling of ENMs, and the application of computational toxicology in assisting the safe use of ENMs. The research objectives and the layout of this thesis are described;

Chapter 2 An inventory of existing toxicity data of metal-based ENMs is established to evaluate relevant data availability and to identify data gaps. The developed database contains 866 data entries on endpoints related to the toxicity of metallic ENMs to algae, yeast, bacteria, protozoa, nematodes, crustacean, and fish;

Chapter 3 In this chapter the development of nano-(Q)SARs is reviewed. The used datasets, constructed models, and underlying mechanisms of ENM uptake and toxicity are discussed;

Chapter 4 Nano-SARs are developed for the categorization of the environmental hazards of metal-based ENMs. Both global nano-SARs across different species and species-specific nano-SARs (for *Danio rerio*, *Daphnia magna*, *Pseudokirchneriella subcapitata*, and *Staphylococcus aureus*) are presented. Possible mechanisms of toxicity are interpreted based on the descriptors used in the models;

Chapter 5 Species sensitivity distributions for metal-based ENMs and relevant HC5 values are obtained. SSDs are developed and compared considering the characteristics of ENMs, the experimental conditions, and different types of endpoints. The most sensitive species and organism groups to certain ENMs are also discussed;

Chapter 6 Based on the presented studies in the thesis and other results in this field, the current knowledge on the use of computational toxicology in assisting the hazard assessment of metallic ENMs is discussed. The development of nano-(Q)SARs and read-across for ENMs, and the development of relevant SSDs are reviewed. Hint messages from the commonly used descriptors in the models are extracted; the toxicity of metal-based ENMs is profiled based on these descriptors. Suggestions and outlook are presented to facilitate the further development of this new frontier.

References

Angel BM, Batley GE, Jarolimek CV, Rogers NJ. The impact of size on the fate and toxicity of nanoparticulate silver in aquatic systems. Chemosphere. 2013, 93:359-65.

Boholm M. The use and meaning of nano in American English: Towards a systematic description. Ampersand. 2016, 3:163-173.

Bondarenko O, Juganson K, Ivask A, Kasemets K, Mortimer M, Kahru A. Toxicity of Ag, CuO and ZnO nanoparticles to selected environmentally relevant test organisms and mammalian cells in vitro: a critical review. Arch Toxicol. 2013, 87:1181-200.

Borm PA, Robbins D, Haubold S, Kuhlbusch T, Fissan H, Donaldson K, Schins R, Stone V, Kreyling W, Lademann J, Krutmann J, Warheit D, Oberdörster E. The potential risks of nanomaterials: A review carried out for ECETOC. Part Fibre Toxicol. 2006, 3:11-46.

Chen G, Li X, Chen J, Zhang YN, Peijnenburg WJ. Comparative study of biodegradability prediction of chemicals using decision trees, functional trees, and logistic regression. Environ Toxicol Chem. 2014, 33:2688-93.

Coll C, Notter D, Gottschalk F, Sun T, Som C, Nowack B. Probabilistic environmental risk assessment of five nanomaterials (nano-TiO_2, nano-Ag, nano-ZnO, CNT, and fullerenes). Nanotoxicology. 2016, 10:436-44.

Commission of the European Communities (CEC), Technical Guidance Document in Support of Commission Directive 93/67/EEC on Risk Assessment for New Notified Substances. Part II, Environmental Risk Assessment, Office for Official Publications of the European Communities, Luxembourg, Luxembourg, 1996.

Dalai S, Pakrashi S, Nirmala MJ, Chaudhri A, Chandrasekaran N, Mandal AB, Mukherjee A. Cytotoxicity of TiO_2 nanoparticles and their detoxification in a freshwater system. Aquat Toxicol. 2013, 138:1-11.

Dams RI, Biswas A, Olesiejuk A, Fernandes T, Christofi N. Silver nanotoxicity using a light-emitting biosensor Pseudomonas putida isolated from a wastewater treatment plant. J Hazard Mater. 2011, 195:68-72.

Donaldson K, Stone V, Clouter A, Renwick L, MacNee W. Ultrafine particles. Occup Environ Med. 2001, 58:211-216.

Dowling A, Clift R, Grobert N, Hutton D, Oliver R, O'Neill O, Pethica J, Pidgeon N, Porritt J, Ryan J, Seaton A, Tendler S, Welland M, Whatmore R. Nanoscience and Nanotechnologies: Opportunities and Uncertainties. Royal Society and Royal Academy of Engineering London, UK, 2004.

Eichert T, Kurtz A, Steiner U, Goldbach HE. Size exclusion limits and lateral heterogeneity of the stomatal foliar uptake pathway for aqueous solutes and water-suspended nanoparticles. Physiol Plant. 2008, 134:151-160.

EU NanoSafety Cluster. Available at https://www.nanosafetycluster.eu/. Accessed on 09.02.2017.

European Chemicals Agency (ECHA). Guidance on information requirements and chemical safety assessment. Chapter R.6: QSARs and grouping of chemicals. 2008a.

European Chemicals Agency (ECHA). Characterization of dose [concentration]-response for environment. Guidance on information requirements and chemical safety assessment. 2008b.

European Chemicals Agency (ECHA). Guidance on Information Requirements and Chemical Safety Assessment, Part B: Hazard Assessment. Version 2.1, European Chemicals Agency: Helsinki, Finland, 2011.

European Commission (EC). Commission recommendation of 18 October 2011 on the definition of nanomaterial (2011/696/EU). 2011.

Fu PP, Xia Q, Hwang HM, Ray PC, Yu H. Mechanisms of nanotoxicity: generation of reactive oxygen species. J Food Drug Anal. 2014, 22:64-75.

Gajewicz A, Schaeublin N, Rasulev B, Hussain S, Leszczynska D, Puzyn T, Leszczynski J. Towards understanding mechanisms governing cytotoxicity of metal oxides nanoparticles: hints from nano-QSAR studies. Nanotoxicology. 2015, 9:313-25.

Galindo TP, Pereira R, Freitas AC, Santos-Rocha TA, Rasteiro MG, Antunes F, Rodrigues D, Soares AM, Gonçalves F, Duarte AC, Lopes I. Toxicity of organic and inorganic nanoparticles to four species of white-rot fungi. Sci Total Environ. 2013, 458-460:290-7.

Gao J, Youn S, Hovsepyan A, Llaneza VL, Wang Y, Bitton G, Bonzongo JC. Dispersion and Toxicity of Selected Manufactured Nanomaterials in Natural River Water Samples: Effects of Water Chemical Composition. Environ Sci Technol. 2009, 43:3322-8.

Garner KL, Suh S, Lenihan HS, Keller AA. Species sensitivity distributions for engineered nanomaterials. Environ Sci Technol. 2015, 49:5753-9.

Gottschalk F, Kost E, Nowack B. Engineered nanomaterials in water and soils: a risk quantification based on probabilistic exposure and effect modeling. Environ Toxicol Chem. 2013, 32:1278-87.

Gottschalk F, Nowack B, Gawlik B. Report on exposure scenarios and release of nanomaterials to the environment. NANEX Work Package 5, EU, FP7 Project Number 247794, 2010.

Gottschalk F, Sonderer T, Scholz RW, Nowack B. Modeled environmental concentrations of engineered nanomaterials (TiO_2, ZnO, Ag, CNT, Fullerenes) for different regions. Environ Sci Technol. 2009, 43:9216-22.

Griffitt RJ, Luo J, Gao J, Bonzongo JC, Barber DS. Effects of particle composition and species on toxicity of metallic nanomaterials in aquatic organisms. Environ Toxicol Chem. 2008, 27:1972-8.

Griffitt RJ, Weil R, Hyndman KA, Denslow ND, Powers K, Taylor D, Barber DS. Exposure to copper nanoparticles causes gill injury and acute lethality in zebrafish (Danio rerio). Environ Sci Technol. 2007, 41:8178-86.

Hall S, Bradley T, Moore JT, Kuykindall T, Minella L. Acute and chronic toxicity of nano-scale TiO_2 particles to freshwater fish, cladocerans, and green algae, and effects of organic and inorganic substrate on TiO_2 toxicity. Nanotoxicology. 2009, 3:91-97.

Heggelund LR, Diez-Ortiz M, Lofts S, Lahive E, Jurkschat K, Wojnarowicz J, Cedergreen N, Spurgeon D, Svendsen C. Soil pH effects on the comparative toxicity of dissolved zinc, non-nano and nano ZnO to the earthworm Eisenia fetida. Nanotoxicology. 2014, 8:559-72.

Heinlaan M, Ivask A, Blinova I, Dubourguier HC, Kahru A. Toxicity of nanosized and bulk ZnO, CuO and TiO_2 to bacteria *Vibrio fischeri* and crustaceans *Daphnia magna* and Thamnocephalus platyurus. Chemosphere, 2008. 71:1308-16.

Hu X, Cook S, Wang P, Hwang HM. In vitro evaluation of cytotoxicity of engineered metal oxide nanoparticles. Sci Total Environ. 2009, 407:3070-2.

Ivask A, Bondarenko O, Jepihhina N, Kahru A. Profiling of the reactive oxygen species-related ecotoxicity of CuO, ZnO, TiO_2, silver and fullerene nanoparticles using a set of recombinant luminescent Escherichia coli strains: differentiating the impact of particles and solubilised metals. Anal Bioanal Chem. 2010, 398:701-16.

Jo HJ, Choi JW, Lee SH, Hong SW. Acute toxicity of Ag and CuO nanoparticle suspensions against Daphnia magna: The importance of their dissolved fraction varying with preparation methods. J Hazard Mater. 2012, 227-228:301-8.

Juganson K, Ivask A, Blinova I, Mortimer M, Kahru A. NanoE-Tox: New and in-depth database concerning ecotoxicity of nanomaterials. Beilstein J Nanotechnol. 2015, 6:1788-804.

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 In Vitro. 2014, 28:600-6.

Kasemets K, Suppi S, Künnis-Beres K, Kahru A. Toxicity of CuO Nanoparticles to Yeast Saccharomyces cerevisiae BY4741 Wild-Type and Its Nine Isogenic Single-Gene Deletion Mutants. Chem Res Toxicol. 2013, 26:356-67.

Keller AA, Lazareva A. Predicted Releases of Engineered Nanomaterials: From Global to Regional to Local. Environ Sci Technol Lett. 2014, 1:65-70

Keller AA, Vosti W, Wang H, Lazareva A. Release of engineered nanomaterials from personal care products throughout their life cycle. J Nanopart Res. 2014, 16:2489.

Klaine SJ, Alvarez PJ, Batley GE, Fernandes TF, Handy RD, Lyon DY, Mahendra S, McLaughlin MJ, Lead JR. Nanomaterials in the environment: behavior, fate, bioavailability, and effects. Environ Toxicol Chem. 2008, 27:1825-51.

Kleandrova VV, Luan F, González-Díaz H, Ruso JM, Melo A, Speck-Planche A, Cordeiro MN. Computational ecotoxicology: simultaneous prediction of ecotoxic effects of nanoparticles under different experimental conditions. Environ Int. 2014, 73:288-94.

Kumar A, Pandey AK, Singh SS, Shanker R, Dhawan A. A flow cytometric method to assess nanoparticle uptake in bacteria. Cytometry A. 2011, 79:707-12.

Kvitek L, Vanickova M, Panácek A, Soukupova J, Dittrich APM, Valentova E, Prucek R, Bancirova M, Milde D, Zboril R. Initial Study on the Toxicity of Silver Nanoparticles (NPs) against Paramecium caudatum. J Phys Chem C. 2009, 113:4296-4300.

Kwok KW, Auffan M, Badireddy AR, Nelson CM, Wiesner MR, Chilkoti A, Liu J, Marinakos SM, Hinton DE. Uptake of silver nanoparticles and toxicity to early life stages of Japanese medaka (Oryzias latipes): effect of coating materials. Aquat Toxicol. 2012, 120-121:59-66.

Lei R, Wu C, Yang B, Ma H, Shi C, Wang Q, Wang Q, Yuan Y, Liao M. Integrated metabolomic analysis of the nano-sized copper particle-induced hepatotoxicity and nephrotoxicity in rats: a rapid in vivo screening method for nanotoxicity. Toxicol Appl Pharmacol. 2008, 232:292-301.

Li M, Czymmek KJ, Huang CP. Responses of Ceriodaphnia dubia to TiO_2 and Al_2O_3 nanoparticles: A dynamic nano-toxicity assessment of energy budget distribution. J Hazard Mater. 2011, 187:502-8.

Liu R, Rallo R, George S, Ji Z, Nair S, Nel AE, Cohen Y. Classification NanoSAR development for cytotoxicity of metal oxide nanoparticles. Small. 2011, 7:1118-26.

Lopes S, Ribeiro F, Wojnarowicz J, Łojkowski W, Jurkschat K, Crossley A, Soares AM, Loureiro S. Zinc oxide nanoparticles toxicity to Daphnia magna: size-dependent effects and dissolution. Environ Toxicol Chem. 2014, 33:190-8.

Lynch I. Water governance challenges presented by nanotechnologies: tracking, identifying and quantifying nanomaterials (the ultimate disparate source) in our waterways. Hydrol Res. 2016, 47:552-568.

Ma H, Bertsch PM, Glenn TC, Kabengi NJ, Williams PL. Toxicity of manufactured zinc oxide nanoparticles in the nematode Caenorhabditis elegans. Environ Toxicol Chem. 2009, 28:1324-30.

Magnuson BA, Jonaitis TS, Card JW. A brief review of the occurrence, use, and safety of food-related nanomaterials. J Food Sci. 2011, 76:R126-33.

Massarsky A, Dupuis L, Taylor J, Eisa-Beygi S, Strek L, Trudeau VL, Moon TW. Assessment of nanosilver toxicity during zebrafish (Danio rerio) development. Chemosphere. 2013, 92:59-66.

Maynard AD, Aitken RJ, Butz T, Colvin V, Donaldson K, Oberdörster G, Philbert MA, Ryan J, Seaton A, Stone V, Tinkle SS, Tran L, Walker NJ, Warheit DB. Safe handling of nanotechnology. Nature. 2006, 444:267-9.

McCracken C, Dutta PK, Waldman WJ. Critical assessment of toxicological effects of ingested nanoparticles. Environ Sci: Nano. 2016, 3:256-282.

Methner M, Hodson L, Dames A, Geraci C. Nanoparticle Emission Assessment Technique (NEAT) for the identification and measurement of potential inhalation exposure to engineered nanomaterials--Part B: Results from 12 field studies. J Occup Environ Hyg. 2010, 7:163-76.

Miralles P, Church TL, Harris AT. Toxicity, Uptake, and Translocation of Engineered Nanomaterials in Vascular plants. Environ Sci Technol. 2012, 46:9224-39.

Modarresi H, Modarress H, Dearden JC. QSPR model of Henry's law constant for a diverse set of organic chemicals based on genetic algorithm-radial basis function network approach. Chemosphere. 2007, 66:2067-76.

Mortimer M, Kasemets K, Kahru A. Toxicity of ZnO and CuO nanoparticles to ciliated protozoa Tetrahymena thermophile. Toxicology. 2010, 269:182-9.

Mulvaney P, Weiss PS. Have Nanoscience and Nanotechnology Delivered? ACS Nano. 2016, 10:7225-6.

Mustad AP, Smeets B, Jeliazkova N, Jeliazkov V, Willighagen EL. Summary of the Spring 2014 NSC Database Survey. Available at: https://figshare.com/articles/Summary_of_the_Spring_2014_NSC_Database_Survey/1195888. Accessed on 07.02.2017.

Nanotechnology Products Database. Available at: http://product.statnano.com/. Accessed on 05.12.2016.

Navarro E, Piccapietra F, Wagner B, Marconi F, Kaegi R, Odzak N, Sigg L, Behra R. Toxicity of silver nanoparticles to Chlamydomonas reinhardtii. Environ Sci Technol. 2008, 42:8959-64.

Nel A, Xia T, Mädler L, Li N. Toxic potential of materials at the nanolevel. Science. 2006, 311:622-7.

Nel AE, Mädler L, Velegol D, Xia T, Hoek EM, Somasundaran P, Klaessig F, Castranova V, Thompson M. Understanding biophysicochemical interactions at the nano-bio interface. Nat Mater. 2009, 8:543-57.

Oberdorster G. Toxicology of ultrafine particles: In vivo studies. Philos Trans R Soc Lond Ser A. 2000, 358:2719-2740.

Oberdörster G, Oberdörster E, Oberdörster J. Nanotoxicology: an emerging discipline evolving from studies of ultrafine particles. Environ Health Perspect. 2005, 113:823-39.

Organization for Economic Cooperation and Development (OECD). Guidance document on the validation of (quantitative) structure activity relationship [(Q)SAR] models. OECD Series on Testing and Assessment No. 69. ENV/JM/MONO(2007)2, Paris, France, 2007.

Park EJ, Yi J, Kim Y, Choi K, Park K. Silver nanoparticles induce cytotoxicity by a Trojan-horse type mechanism. Toxicol In Vitro. 2010, 24:872-8.

Posthuma L, Suter II GW, Traas TP. Species sensitivity distributions in ecotoxicology. Lewis Publishers, CRC Press, Boca Raton, FL, USA, 2002.

Project on Emerging Nanotechnologies. Consumer Products Inventory. Available at http://www.nanotechproject.org/cpi. Accessed on 23.11.2016.

Puzyn T, Leszczynska D, Leszczynski J. Toward the development of "nano-QSARs": advances and challenges. Small. 2009, 5:2494-509.

Puzyn T, Rasulev B, Gajewicz A, Hu X, Dasari TP, Michalkova A, Hwang HM, Toropov A, Leszczynska D, Leszczynski J. Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles. Nat Nanotechnol. 2011, 6:175-8.

Raies AB, Bajic VB. 2016. In silico toxicology: computational methods for the prediction of chemical toxicity. Wiley Interdiscip Rev Comput Mol Sci. 2016, 6:147-172.

Reisfeld B, Mayeno AN. What is computational toxicology? Methods Mol Biol. 2012, 929:3-7.

Russell WMS, Burch RL. The Principles of Humane Experimental Technique, Methuen, London, 1959.

Sahoo SK, Parveen S, Panda JJ. The present and future of nanotechnology in human health care. Nanomedicine. 2007, 3:20-31.

Samberg ME, Orndorff PE, Monteiro-Riviere NA. Antibacterial efficacy of silver nanoparticles of different sizes, surface conditions and synthesis methods. Nanotoxicology. 2011, 5:244-53.

Sealy C. 'Safe-By-Design' Nanoparticles Show Reduced Toxicity. Nano Today. 2011, 6:113-114.

Shaw SY, Westly EC, Pittet MJ, Subramanian A, Schreiber SL, Weissleder R. Perturbational profiling of nanomaterial biologic activity. Proc Natl Acad Sci U S A. 2008, 105:7387-92.

Shi JP, Ma CY, Xu B, Zhang HW, Yu CP. Effect of light on toxicity of nanosilver to Tetrahymena pyriformis. Environ Toxicol Chem. 2012, 31:1630-8.

Singh KP, Gupta S, Kumar A, Mohan D. Multispecies QSAR modeling for predicting the aquatic toxicity of diverse organic chemicals for regulatory toxicology. Chem Res Toxicol. 2014, 27:741-53.

Sizochenko N, Leszczynski J. Review of current and emerging approaches for quantitative nanostructure–activity relationship modeling: the case of inorganic nanoparticles. J Nanotoxicol Nanomed. 2016, 1:1-16.

StatNano. Available at http://statnano.com/. Accessed on 09.02.2017.

Sun TY, Gottschalk F, Hungerbühler K, Nowack B. Comprehensive probabilistic modelling of environmental emissions of engineered nanomaterials. Environ Pollut. 2014, 185:69-76.

Tantra R, Oksel C, Puzyn T, Wang J, Robinson KN, Wang XZ, Ma CY, Wilkins T. Nano(Q)SAR: Challenges, pitfalls and perspectives. Nanotoxicology. 2015, 9:636-42.

Thakkar KN, Mhatre SS, Parikh RY. Biological synthesis of metallic nanoparticles. Nanomedicine. 2010, 6:257-62.

Tyne W, Lofts S, Spurgeon DJ, Jurkschat K, Svendsen C. A new medium for Caenorhabditis elegans toxicology and nanotoxicology studies designed to better reflect natural soil solution conditions. Environ Toxicol Chem. 2013, 32:1711-7.

Valsami-Jones E, Lynch I. NANOSAFETY. How safe are nanomaterials? Science. 2015, 350:388-9.

Vance ME, Kuiken T, Vejerano EP, McGinnis SP, Hochella MF Jr, Rejeski D, Hull MS. Nanotechnology in the real world: Redeveloping the nanomaterial consumer products inventory. Beilstein J Nanotechnol. 2015, 6:1769-80.

von Moos N, Slaveykova VI. Oxidative stress induced by inorganic nanoparticles in bacteria and aquatic microalgae-state of the art and knowledge gaps. Nanotoxicology. 2014, 8:605-30.

Wang L, Nagesha DK, Selvarasah S, Dokmeci MR, Carrier RL. Toxicity of CdSe Nanoparticles in Caco-2 Cell Cultures. J Nanobiotechnology. 2008, 6:11.

Wang Z, Chen J, Li X, Shao J, Peijnenburg WJ. Aquatic toxicity of nanosilver colloids to different trophic organisms: contributions of particles and free silver ion. Environ Toxicol Chem. 2012, 31:2408-13.

Winkler DA. Recent advances, and unresolved issues, in the application of computational modelling to the prediction of the biological effects of nanomaterials. Toxicol Appl Pharmacol. 2016, 299:96-100.

Xiao Y, Vijver MG, Chen G, Peijnenburg WJ. Toxicity and accumulation of Cu and ZnO nanoparticles in Daphnia magna. Environ Sci Technol. 2015, 49:4657-64.

Yang X, Gondikas AP, Marinakos SM, Auffan M, Liu J, Hsu-Kim H, Meyer JN. Mechanism of Silver Nanoparticle Toxicity Is Dependent on Dissolved Silver and Surface Coating in Caenorhabditis elegans. Environ Sci Technol. 2012, 46:1119-27.

Zhang H, Ji Z, Xia T, Meng H, Low-Kam C, Liu R, Pokhrel S, Lin S, Wang X, Liao YP, Wang M, Li L, Rallo R, Damoiseaux R, Telesca D, Mädler L, Cohen Y, Zink JI, Nel AE. Use of metal oxide nanoparticle band gap to develop a predictive paradigm for oxidative stress and acute pulmonary inflammation. ACS Nano. 2012, 6:4349-68.

Zhou H, Mu Q, Gao N, Liu A, Xing Y, Gao S, Zhang Q, Qu G, Chen Y, Liu G, Zhang B, Yan B. A nano-combinatorial library strategy for the discovery of nanotubes with reduced protein-binding, cytotoxicity, and immune response. Nano Lett. 2008, 8:859-65.

Zhu X, Tian S, Cai Z. Toxicity assessment of iron oxide nanoparticles in zebrafish (Danio rerio) early life stages. PLoS One. 2012, 7:e46286.

CHAPTER 2

SUMMARY AND ANALYSIS OF THE CURRENTLY EXISTING LITERATURE DATA ON METAL-BASED NANOPARTICLES PUBLISHED FOR SELECTED AQUATIC ORGANISMS: APPLICABILITY FOR TOXICITY PREDICTION BY (Q)SARS

Chen G, Vijver MG, Peijnenburg WJGM

Published in *Alternatives to laboratory animals*. 2015, 43(4): 221-40

Abstract

This review establishes an inventory of existing toxicity data on nanoparticles (NPs) with the purpose of developing (Quantitative) Structure–Activity Relationships for NPs (nano-(Q)SARs) and also of maximizing the use of scientific sources for NP risk assessment. From a data search carried out on 27 February 2014, a total of 910 publications were retrieved from the Web of Science™ Core Collection, and a database comprising 886 records of toxicity endpoints was built based on these publications. The test organisms mainly comprised bacteria, algae, yeast, protozoa, nematode, crustacean, and fish. The NPs consisted mainly of metals, metal oxides, nanocomposites, and quantum dots. The data were analyzed further, in order to: i) categories each toxicity endpoint and the biological effects triggered by the NPs; ii) survey the characterization of the NPs used; and iii) assess whether the data were suitable for nano-(Q)SAR development. Despite the efforts of numerous scientific programmes on nanomaterial safety and design, our study concluded that lack of data consistency prevents the use of experimental data in developing and validating nano-(Q)SARs. Finally, an outlook on the future of nano-(Q)SAR development is provided.

Key words: ecotoxicity, metal-based, models, nanoparticle, (quantitative) structure–activity relationship

2.1 Introduction

Tremendous advances in the utility of synthetic nanoparticles (NPs) have raised global concerns about potential nano-specific effects on ecosystems. The likelihood of NPs triggering negative impacts on ecosystems, as well as on human health, has already been addressed by various studies (Gajewicz et al., 2012; Ivask et al., 2014; Schrand et al., 2010). This likelihood necessitates a comprehensive risk assessment of NPs, to determine whether their benefits outweigh the risks, before initiating large-scale production. Such a task, however, is prevented by insufficient scientific information, as evident from the observation that the number of studies investigating the harmful effects of NPs severely lags behind the rapid growth of nanotechnology (Bondarenko et al., 2013; Kahru and Ivask, 2013). In addition, the exponential increase in the number and variety of NPs makes it impossible to test every newly-synthesized NP, taking into account the high study-cost, the time-consuming nature of toxicity testing and the Three Rs (replacement, reduction, and refinement) concept governing animal use (Russell and Burch, 1959). If the use of alternative non-animal approaches was maximized for testing, 1.9 million fewer animals would be required (Gajewicz et al., 2012). Therefore, reliable protocols for *in silico* screening of the effects of NPs are required for adequate NP risk assessment (Cumming et al., 2013; OECD Quantitative Structure–Activity Relationships Project). Meanwhile, the mission of safe-by-design for nanotechnology (Maynard et al., 2006), which was aimed at designing biologically and environmentally benign NPs, has also driven the need for predicting the toxicity of NPs from their pristine (i.e. unmodified) structures. Thus, recently, there have been many attempts to predict the toxicity of NPs based on computational methods.

Following its successful application in formalizing relationships between structural characteristics and biological effects (Altschuh et al., 1999; Arnot and Gobas, 2006; Pavan and Worth, 2006; Pavan et al., 2006; Chen et al., 2014), the (Quantitative) Structure–Analysis Relationship ((Q)SAR) approach offers a rapid way of filling data gaps caused by limited availability, or the absence, of experimental information. Attempts have also been made to use experimentally-obtained data to link the physical-chemical characteristics of NPs to their cellular uptake, cytotoxicity (Epa et al., 2012; Ehret et al., 2014; Fourches et al., 2010; Gajewicz et al., 2015; Ghorbanzadeh et al., 2012; Kar et al., 2014a; Liu et al., 2011; 2013; Luan et al., 2014; Singh and Gupta, 2014; Sizochenko et al., 2014; Toropov et al., 2013), and ecotoxicity (Kar et al., 2014b; Kleandrova et al., 2014; Pathakoti et al., 2014; Puzyn et al., 2011; Singh and Gupta, 2014; Sizochenko et al., 2014; Toropov et al., 2012). The nano-(Q)SARs reported and the data sets used are summarized in Table 2.1. As can be seen from this table, despite the fact that intensive research is being carried out on NP-related toxicity, the nano-(Q)SARs developed so far have mainly employed toxicity information from a limited number of studies on a restricted number of classes of NPs, and

used data generated under consistently similar conditions. To improve the development of nano-(Q)SARs, with the ultimate goal of employing nano-(Q)SARs as alternative *in silico* screening methods in toxicity testing, it is essential that all published nanotoxicity data is summarized and organized into potentially useful data for modeling researchers. Meanwhile, NP-related regulatory frameworks also require the gathering of nanotoxicity information to enable the optimal use of the existing scientific sources. For instance, according to EU Directive 93/67/EEC, it is preferable that the classification of chemical hazard to aquatic organisms is based on toxicity data from at least three standard test organisms (i.e. algae, crustacean, and fish), with hazard initially determined by the lowest median L(E)C50 value (the chemical concentration found to cause 50% death or effect of interest) of the species tested: if the L(E)C50 value is < 1 mg/L, the compound is considered very toxic to aquatic organisms; if the L(E)C50 value is 1–10 mg/L, the compound is considered toxic to aquatic organisms; if the L(E)C50 value is 10–100 mg/L, the compound is considered harmful to aquatic organisms; if the L(E)C50 value is > 100 mg/L, the compound is not classified as being toxic or harmful (Ivask et al., 2014; Commission of the European Communities, 1996). Blaise et al. (2008) and Sanderson et al. (2003) have subsequently extended this classification scheme by adding one more category: L(E)C50 value < 0.1 mg/L corresponds to compounds that are extremely toxic to aquatic organisms (Ivask et al., 2014; Kahru and Dubourguier, 2010).

Two kinds of data are essential for developing predictive (Q)SAR models: data that characterize the physico-chemical properties of groups of pristine NPs (NP descriptors), and data that describe the relevant biological effects of NPs on test organisms (toxicity endpoints), including a detailed description of the experimental conditions, or of the test protocols used (Ivask et al., 2014). It is also generally acknowledged that NP dynamics in the test medium (e.g. aggregation, agglomeration, dissolution) greatly impact their toxicity (El Badawy et al., 2010; Tiede et al., 2009); the applicability of (Q)SARs would be broadened, if such transformations could be incorporated into the data input when linking NP characteristics to toxicity. Nevertheless, the few successful efforts that have been made to develop nano-(Q)SARs were restricted to correlating the characteristics of pristine NPs with NP toxicity. Puzyn et al. (2011) modeled the toxicity of 17 metal oxide NPs in *Escherichia coli* (*E. coli*) employing only one descriptor ΔH_{Me+} (the enthalpy of formation of a gaseous cation having the same oxidation state as that in the metal oxide structure). Based on the same data set, Singh and Gupta (2014) recently built a nano-(Q)SAR model with three descriptors: oxygen percentage, molar refractivity, and polar surface area. Other models were also constructed by using the same data set (Table 2.1), solely on the basis of the descriptors of pristine NPs (Kar et al., 2014b; Toropov et al., 2012; Sizochenko et al., 2014). Metal electronegativity (χ) and the charge of the metal cation corresponding to a given oxide (χ_{ox}) were also employed to predict photo-induced toxicity of 17 oxide NPs to

E. coli (Pathakoti et al., 2014). Those studies indicated that it is indeed possible to predict NP toxicity based on the characteristics of pristine NPs, which would benefit the development of *in silico* screening protocols as an alternative to experimental assays, as well as complying with the 'safe-by-design' initiative for nanotechnology.

2.2 Methods

As a first step toward the development of nano-(Q)SARs, and based on a Web of Science™ Core Collection bibliometric data search, we established an inventory of toxicity data of metal-based NPs that are widely used in a variety of applications (Schrand et al., 2010). Information on NP characterization, if it was associated with the reported toxicity data, was also included. The focus of organisms were based on the studies of Ivask et al. (2014), Bondarenko et al. (2013), and Kahru and Dubourguier (2010). These are mainly algae, yeast, bacteria, protozoa, crustacean, nematodes, and fish. The findings were evaluated in the light of nano-(Q)SAR development and were based on the characteristics of pristine NPs. The toxicity endpoints reported in the literature, and hence included in the database, mainly consisted of the lethal concentration (LC), the effect concentration (EC, or IC when the effect refers to inhibition), the lowest observed effect concentration (LOEC), the no observed effect concentration (NOEC), the minimum bactericidal concentration (MBC), and the minimum inhibitory concentration (MIC; more commonly used in antimicrobial assays). Information on these common toxicity endpoints was extracted from the retrieved publications. The test species and metal-based NPs covered in this review provide an overview of the database. For (Q)SAR modeling purposes, further analysis focused on the numbers of different toxicity endpoints, the type of biological effects induced by the NPs, data availability (i.e. the amount of accessible toxicity data), and also the characterization of the NPs provided.

2.2.1 Bibliometric data search

To access the experimental information available, a bibliometric data search was performed on 27 February 2014 by using the Advanced Search features in the Web of Science Core Collection. To ensure that the data search, and subsequent analysis, covered a broad range of test species, different hierarchies of organisms were selected based on the studies by Ivask et al. (2014), Bondarenko et al. (2013), and Kahru and Dubourguier (2010). The test organisms analyzed in this review mainly comprised: bacteria, algae, yeast, protozoa, nematode, crustacean, and fish. The NPs selected for this review included a variety of metal-based NPs, based on an empirical analysis of existing nanotoxicity-related

publications. The test species and metal-based NPs were subsequently identified by using (truncated) search terms (i.e. key words) as given in Tables S2.1 and S2.2 in the Supplemental Information. NPs with no search records evident after a preliminary search in the Web of Science Core Collection were excluded. To eliminate redundant records, the data search was restricted by two conditions: a) the research area had to be toxicology, or the topic contain "*toxicity" and "effect*" but not "function*", "synthesis", "label" or "agent", to exclude studies on related applications; and b) either title or abstract had to contain "nano" or "quantum". The language and type of document were restricted, respectively, to "English" and "article". Finally, 23 different kinds of metal-based NPs were included in the study: silver (Ag), aluminium (Al), gold (Au), bismuth (Bi), cadmium (Cd), cerium (Ce), cobalt (Co), chromium (Cr), copper (Cu), iron (Fe), indium (In), lanthanum (La), manganese (Mn), molybdenum (Mo), nickel (Ni), platinum (Pt), antimony (Sb), selenium (Se), silicon (Si), titanium (Ti), vanadium (V), zinc (Zn), and zirconium (Zr).

2.3 Results and Discussion

2.3.1 Overall analysis of the NP-related studies

A total of 982 papers were retrieved, according to the data search refined by condition (a). Most papers featured bacteria as the test organisms and with silver NPs as the metal-based NPs (Table S2.3). After assigning condition (b), a total of 910 papers were obtained. A detailed analysis showed that 406 papers described studies on bacteria, 245 on fish, 193 on crustacean, 134 algae, 102 yeast, 43 nematodes, and 17 protozoa (Figure 2.1). With regard to the metal-based NPs, 383 papers were related to the toxicity induced by silver NPs, followed by 238 on titanium, 139 on copper and 137 on zinc NPs (Figure 2.2 and Table S2.4). Of the 910 papers, 45 dealt specifically with quantum dots, with either the title or abstract containing the key word "quantum dot".

2.3.2 Analysis of toxicity endpoints

Of the 910 papers retrieved, a manual selection was subsequently carried out to screen data related to the aforementioned toxicity endpoints (LC, EC, LOEC, NOEC, MBC, and MIC). A database with 886 records of the toxicity endpoints was obtained and summarized in a Microsoft Excel® spreadsheet (see Supplementary Information, available on the *ATLA* website www.atla.org.uk). The original data were presented according to the following features:

Table 2.1. Summary of experimental data used for the reported nano-(Q)SARs

Reference of dataset	Nanomaterials (NMs) covered	Biologic effects	Type of organisms or cells	Reported nano-(Q)SARs
Weissleder et al., 2005	146 NMs with $(Fe_2O_3)_m(Fe_3O_4)_n$ core but different surface modifiers	Cellular uptake	Pancreatic cancer cells; macrophage cell line; resting primary human macrophages; activated primary human macrophages; human umbilical vein endothelial cells	Chau and Yap, 2012 Kar et al., 2014a Epa et al., 2012 Fourches et al., 2010 Ghorbanzadeh et al., 2012 Singh and Gupta, 2014
Puzyn et al., 2011 (data partly from Hu et al., 2009)	17 metal oxide NMs	Ecotoxicity	Escherichia coli	Kar et al., 2014b Puzyn et al., 2011 Singh and Gupta, 2014 Sizochenko et al., 2014 Toropov et al., 2012
Shaw et al., 2008	48 $(Fe_2O_3)_m(Fe_3O_4)_n$ core based NMs and two quantum dots	Cytotoxicity	Endothelial cells; vascular smooth muscle cells; hepatocytes; murine RAW 264.7 leukemic monocyte/macrophage cells	Epa et al., 2012 Singh and Gupta, 2014 Ehret et al., 2014 Fourches et al., 2010
Gajewicz et al., 2015	18 metal oxide NMs	Cytotoxicity	Human keratinocyte cells	Gajewicz et al., 2015 Sizochenko et al., 2014
Pathakoti et al., 2014	17 metal oxide NMs	Ecotoxicity	Escherichia coli	Pathakoti et al., 2014
Liu et al., 2011	9 metal oxide NMs	Cytotoxicity	Transformed bronchial epithelial cells	Liu et al., 2011
Zhang et al., 2012	24 metal oxide NMs	Cytotoxicity	Human bronchial epithelial cells; rat alveolar macrophage cells	Liu et al., 2013
Multi data sources				Kleandrova et al., 2014 Luan et al., 2014

(i) References, including first author, publication year, journal and title of the publication;

(ii) Organism details, i.e. the categorical group, the species, bacterial strain or life-stage used;

(iii) Experimental conditions, including the duration of exposure, type of light exposure and emittance of light (for phototoxicological studies), media composition, and pH (when the experiments were based on standardized tests, e.g. OECD guidelines, the name of the test was given instead);

(iv) Toxicity endpoints, as described by the biological effect addressed, type of endpoint, experimental value of toxicity endpoint, and the unit used; and

(v) NP characterization, consisting of the type of NP, core, size, coating, purity, crystallinity, surface area, surface charge, shape, and zeta potential.

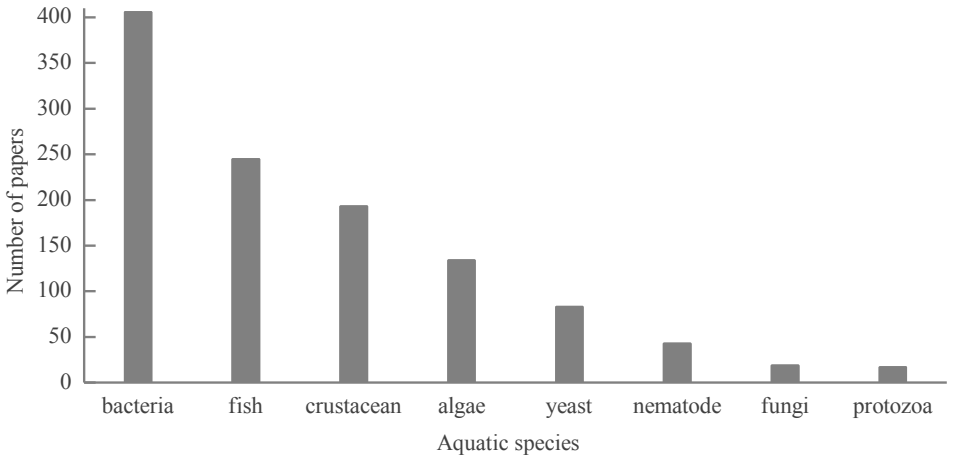


Figure 2.1. Number of retrieved papers on the organisms in the Web of Science™ Core Collection. Data search was performed in the Web of Science™ Core Collection on 27 February, 2014; key words used characterizing tested organisms were listed in Table S2.1. The organism-wise analysis based on 910 retrieved publications from the data search indicates that bacteria is the most generally studied organism for testing nanotoxicity, followed by fish, crustacean, algae, yeast, nematode, and protozoa.

The database covered 62 species (55 species that comprised the original seven test organisms, plus additional data on seven species) and 29 kinds of metal-based NPs in total.

It included 20 species of bacteria, 12 species of algae, 5 species of yeast, 4 species of protozoa, 2 species of nematodes, 7 species of crustacean, and 5 species of fish. These 55 species were found to be related to 866 toxicity endpoints presented in the database. The main journals where these toxicity endpoints were published were: *Nanotoxicology* (128 records), *Environmental Toxicology and Chemistry* (94 records), *Environmental Science and Technology* (91 records), *Chemosphere* (67 records), and *Science of the Total Environment* (61 records). To highlight the main points of the database, test species with at least six records (in total 28 species with 802 records) are shown in Figure 2.3. Toxicity endpoints in other organisms that were studied simultaneously in the retrieved publications were also collected and included in the Supplementary Information. As Figure 2.3 shows, most of the NPs were metals, metal oxides, nanocomposites, and quantum dots. With regard to toxicity endpoints, E(I)C was the most recorded (accounting for 444 records), followed by LC (with 187 records, two of which lethal dose), MIC (112 records), NOEC (50 records), LOEC (44 records), and MBC (49 records). The numbers of toxicity endpoints involving certain species and specific metal-based NPs were also analyzed, as shown in Figure 2.3.

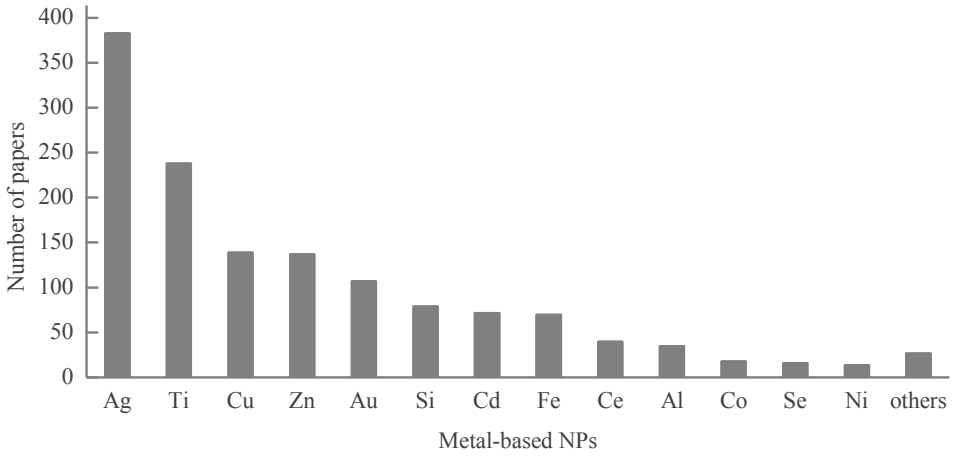


Figure 2.2. Number of retrieved papers on metal-based NPs in the Web of Science™ Core Collection. A comparison on number of publications concerning the toxicity studies of different metal-based NPs. NPs with less than ten papers are shown in the group "others", namely Pt, Cr, In, Zr, Bi, La, Mn, Mo, Sb, and V NPs. Data search was performed on 27 February, 2014 in the Web of Science™ Core Collection, key words used characterizing the NPs were given in Table S2.2. It can be seen that Ag NPs attracted the most research attention among the metal-based NPs.

The analysis indicated that Ag NPs were the most widely studied NPs (with a total of 332 records of endpoints), with a particular focus on two bacteria, *Staphylococcus aureus* (*S. aureus*) and *E. coli*, and a crustacean, *Daphnia magna* (*D. magna*). Meanwhile, more than average attention was also paid to TiO_2 (126) and ZnO (109) NPs. As for the test organisms most often used, *D. magna*, *E. coli* (a bacterium), and *Pseudokirchneriella subcapitata* (an alga) were the dominant species in the database, with 173, 139, and 106 toxicity records, respectively. They were followed by *S. aureus* (49 toxicity records), *Vibrio fischeri* (a bacterium; 47 records), and *Danio rerio* (the zebrafish; 44 records). Given the numbers of available records, these data are potentially useful for nano-(Q)SAR modeling, but care should be taken regarding data consistency. If we take the endpoints in *D. magna* as an example, the 173 records retrieved on this water flea consisted of 51 values for LC50, 67 values for EC50, 19 values for NOEC, 13 values for LOEC, and 23 others (e.g. LC10, LC20, EC10, EC20, etc.). In addition, there was further variation with regard to the various biological effects that were assessed and the duration of the exposure of the organisms to the NPs.

The NP-induced biological effects and relevant toxicity endpoints are shown in Table 2.2. The biological effects commonly investigated include: mortality, cell viability inhibition, growth inhibition, immobilization, luminescence inhibition, malformation, and reproduction inhibition. Mortality and growth inhibition are two significant indices that are generally applied in ecotoxicity assays; the rest of the endpoints are used as appropriate on different groups of organisms. Unsurprisingly, as the standard test organisms in OECD guidelines, algae, bacteria, crustacean, and fish are paid relatively more research attention. Immobilization, for instance, is an important factor that is often used to characterize the effects of NPs on crustacean. Inhibition of reproduction is another commonly studied endpoint. Meanwhile, luminescence inhibition is examined only with bacteria. With regard to the issue of data availability of toxicity endpoints for nano-(Q)SAR modeling, the EC50 (growth inhibition) to algae accounts for 91 records. Concerning bacteria, 110 MIC data records and 86 EC50 (luminescence inhibition) values were retrieved. For crustacean, 82 LC50 (mortality) and 59 EC50 (immobilization) values were found. For fish, 44 LC50 (mortality) values were obtained.

Besides the test species used and the biological effects and toxicity endpoints measured, the diversity of metal-based NPs in a data set is also of major importance for the development of nano-(Q)SARs. In this context, two issues stand out: first, regardless of the number of records available in a data set with the same toxicity endpoint, the data set should cover different NPs in order to be potentially modeled against NP properties; second, the NPs of interest should share a degree of structural similarity in order to be grouped and described in terms of descriptors suitable for modeling.

Table 2.2. Biological effects to eight groups of organisms and corresponding toxicity endpoints

	Algae	**Bacteria**	**Crustacean**	**Fish**	**Nematode**	**Protozoa**	**Yeast**
Mortality	LC50 (1)	LC50 (23) MBC (49)	LC5 (1) LC10 (3) LC15 (1) LC20 (6) LC25 (1) LC50 (82) LOEC (10) NOEC (16)	LC50 (44) LD50 (2) LOEC (4)	LC50 (7)	LC50 (8)	
Cell viability inhibition	EC50 (2)	EC50 (4)				EC50 (4)	EC50 (10)
Growth inhibition	EC10 (21) EC20 (19) EC30 (2) EC50 (91) EC100 (2) LOEC (6) NOEC (15)	EC50 (13) MIC (110) LOEC (15) NOEC (6)	EC50 (1)	EC50 (2)	EC50 (14)	EC50 (6) EC50 (1)	EC20 (5) EC50 (12)
Immobilization			EC10 (10) EC50 (59) NOEC (3)		EC50 (3)		
Luminescence inhibition		EC10 (2) EC20 (15) EC50 (86) MIC (2) LOEC (2) NOEC (2)					
Malformation				EC50 (10)			
Reproduction inhibition	EC50 (2)		EC10 (5) EC20 (4) EC50 (13) NOEC (4) LOEC (4)		EC50 (2)		
Others[a]		EC50 (9) NOEC (2)	EC50 (2) LOEC (1)	EC50 (4) NOEC (1)			

[a] e.g. colony formation inhibition, feeding inhibition, fertilization inhibition, hatching delay; EC - effect concentration; LC - lethal concentration; LD - lethal dose; LOEC - lowest observed effect concentration; MBC - minimum bactericidal concentration; MIC - minimum inhibitory concentration; NOEC - no observed effect concentration. The number of record is provided in parentheses. The data shown are summarized from the database provided in the Supplementary Information.

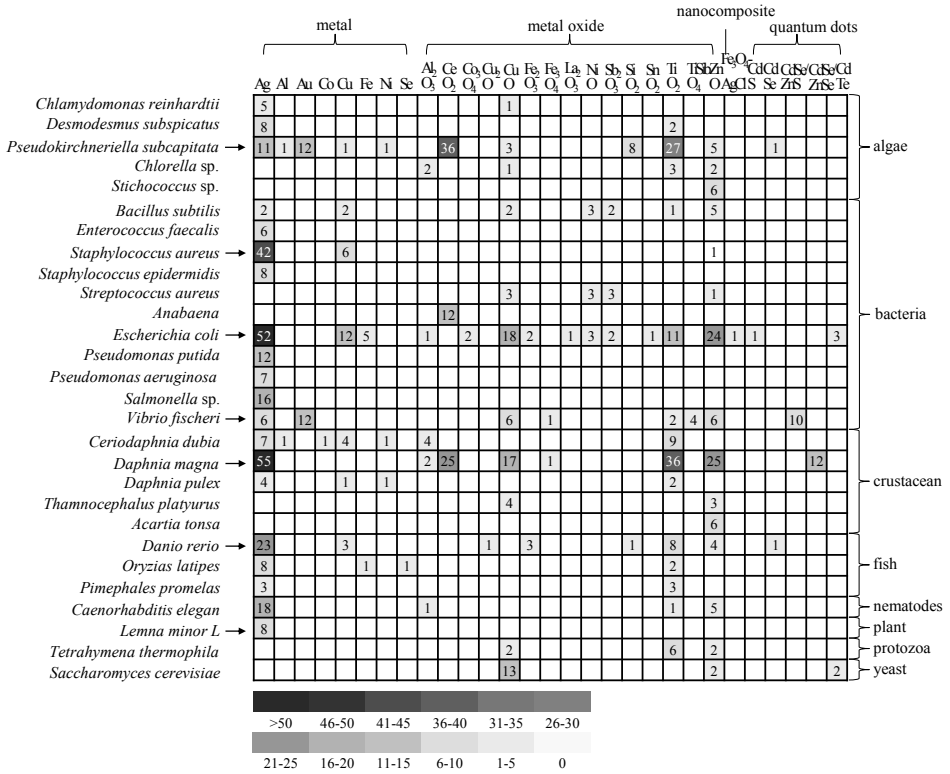


Figure 2.3. Overview of the database regarding various metal-based NPs, tested species, and numbers of records of the toxicity endpoints. This analysis is based on the database provided in the Supplementary Information (Excel spreadsheet), which was retrieved on 27 February, 2014. To illustrate the main information in the database, test species with fewer than six records are not shown (in total, 34 species with 84 records). Details of the references, test organisms, experimental conditions, NP properties, and toxicity endpoints are also listed in the Supplementary Information.

A diversity analysis of the metal-based NPs was performed on the toxicity data of the six species mentioned above (Table S2.5): *D. magna* (173 records), *E. coli* (139 records), *P. subcapitata* (106 records), *S. aureus* (49 records), *V. fischeri* (47 records), and *D. rerio* (44 records). LC50 values for *D. magna*, *E. coli*, and *D. rerio* are reported for eight NPs. Meanwhile, growth inhibition to *P. subcapitata* (EC50) was reported for ten NPs. These endpoints could possibly be considered for building nano-(Q)SAR models. Moreover, median L(E)C50 values of the metal-based NPs to the organisms were analyzed for three

purposes, as shown in Figure 2.4 (see details in Table S2.6). The first purpose was to identify potentially hazardous NPs, with the aim of focusing modeling for the most hazardous NPs. Adhering to EU Directive 93/67/EEC (Commission of the European Communities, 1996; Kahru et al., 2010), and studies of Blaise et al. (2008) and Sanderson et al. (2003), the metal-based NPs were classified in five hazard categories, as shown in Figure 2.4. Based on this distinction, Ag and Cu NPs needed to be classified as 'very toxic' to aquatic organisms. The 'toxic' category included Ce, Co, Ni, Se, Ti, and Zn NPs, while Al, Au, and Fe NPs were considered to be 'harmful'. Data on La, Sb, and Sn NPs were totally absent, which might be due to less research interest and/or missing information in the data search.

We also compared toxicity data on specific NPs or organisms in order to identify the most toxic NP for each organism, or the most sensitive organism to a certain NP. For instance, among NPs, those that are Ag-based have the lowest median L(E)C50 values (most toxic) to algae (0.1 mg metal/L), crustacean (0.01 mg metal/L), and nematodes (2.85 mg metal/L). Crustaceans are more sensitive to Ag (0.01 mg metal/L) and Cu NPs (0.61 mg metal/L), as compared to other test organisms. In order to develop nano-(Q)SARs, a range of values for the toxicity data for a given species is also needed, to permit modeling against NP properties. According to the study by Song et al. (2011), a feasible strategy might be to model a large variation of toxicity values of certain NPs against the ecological traits of the organisms. The analysis of the data retrieved shows that our toxicity data have a large variation of toxicity values for both metal-based NPs and for the test organisms used, thus potentially allowing the development of nano-(Q)SARs for a limited number of endpoints or for a limited number of species.

2.3.3 Characterization of the metal-based NPs

In addition to data availability on toxicity endpoints, NP characterization in the form of measured and/or calculated NP properties, also plays an essential role in the development of nano-(Q)SARs. Figure 2.5a shows the frequency distribution of the measured NP properties in the data retrieved. The NP properties analyzed included: zeta potential, surface charge and surface area, size, shape, purity, crystallinity, and coating. Our results show that the size (primary) of metal-based NPs was generally provided (847 records), followed by the zeta potential (316 records), surface area (224 records), coating (117 records), and purity (87 records). Only a limited number of studies offered information about the shape (67 records), crystallinity (57 records), and surface charge (three records). It is worth noting that some of the data on these properties are hardly suited for the purpose of nano-(Q)SAR studies. For instance, NP size is on occasions given as a range between 20–60 nm or < 100 nm (Gladisa et al., 2010; Jo et al., 2012), which is not precise enough for developing models. Thus, for

the purpose of nano-(Q)SAR development, data availability on measured NP properties is even more limited than that reported. As shown in Figure 2.5b, after an analysis of the number of measured properties for a certain NP in a publication, it is clear that relatively few NP features are usually investigated. Most of the published toxicity endpoints contain one or two NP properties, and only 5.3% of the assembled records contain more than three NP properties. Thus, according to our analysis of the availability of NP properties, the development of nano-(Q)SAR models simply on the basis of the reported experimentally determined descriptors would, at this time, be a challenging task.

2.3.4 Comparing the results to other databases

Recently, Oksel et al. (2015) reviewed literature data that was suitable for developing nano-QSARs. They summarized data sets from eight studies concerning both the toxicity endpoints of interest and relevant NP characterization; the data are presented in the supplementary information of the original publication. In addition to the experimental data published in the scientific literature, some online databases are being developed under various projects and can be used as sources for retrieving experimental data, as described in the Summary of the Spring 2014 NSC Database Survey (2014). For example, the Nanomaterial-Biological Interactions (NBI) Knowledgebase (http://nbi.oregonstate.edu/) is an online database that also contains information on the toxicity of nanomaterials. It includes data on NP toxicity to zebrafish embryos, based on an indicator that integrates observed mortality, immobilization, and malformation. The distribution of the types of metal-based NPs and NP characterization in the NBI was analyzed and compared to that in our database (Figure S2.1 and S2.2). The results show that both databases contain toxicity data of metal and metal oxide NPs and nanocomposites, of which metal oxide NPs are the dominant group, followed by metal NPs. With regard to NP characterization, except for NP primary size, the two databases emphasize different properties (see Figure S2.2): the NBI database provides more data concerning the functional group, shape, purity, and surface charge; our database has relatively more records of zeta potential and surface area. This comparison reveals differences in the NP properties measured for characterization, but it also suggests a high similarity between the main types of NPs presented. We thus conclude that our review of the literature on NP toxicity should be considered representative of the actual situation with regard to data availability and data quality.

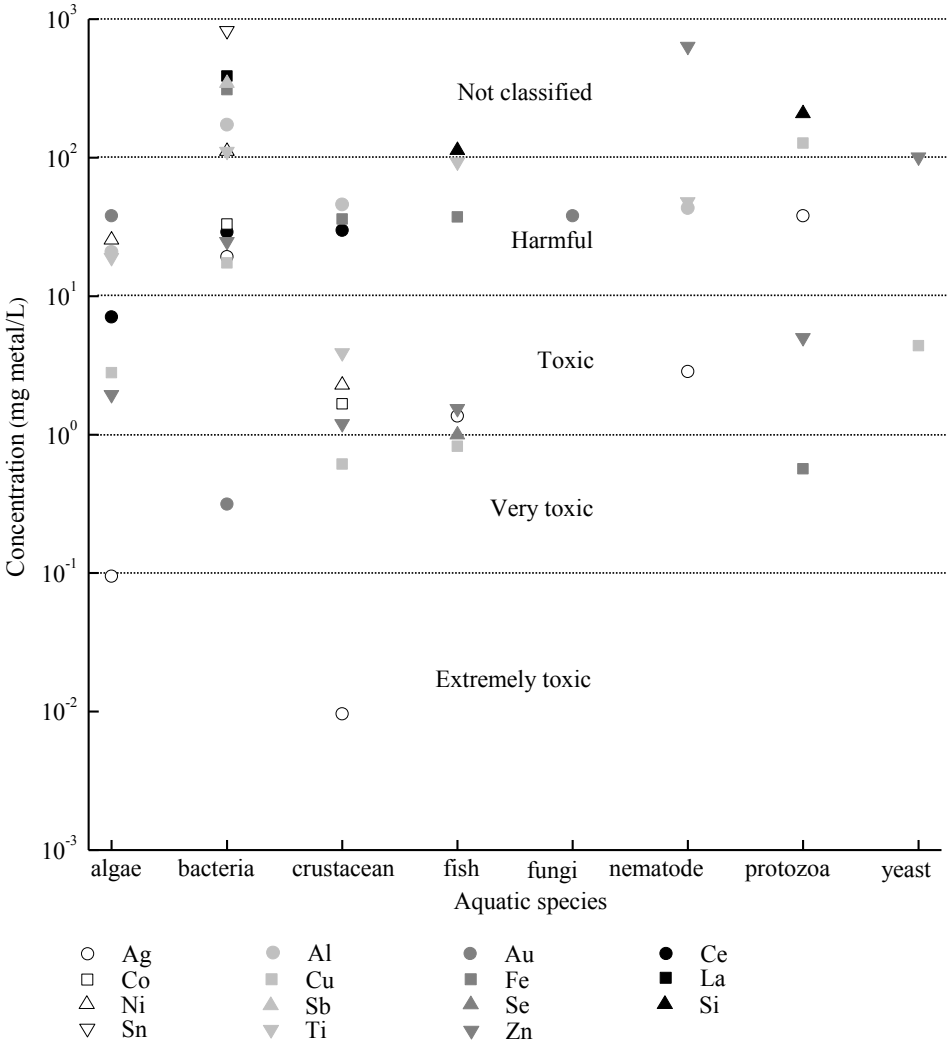


Figure 2.4. Median L(E)C50 values of metal-based NPs to organisms. The classification of hazard categories for the NPs adheres to the EU-Directive 93/67/EEC (Commission of the European Communities, 1996), and the studies of Blaise et al. (2008) and Sanderson et al. (2003). NPs are grouped as not classified, harmful, toxic, very toxic, and extremely toxic to aquatic organisms based on the lowest median L(E)C50 value for the organisms (algae, crustacean, and fish): < 0.1 mg/L = extremely toxic to aquatic organisms; 0.1–1 mg/L = very toxic to aquatic organisms; 1–10 mg/L = toxic to aquatic organisms; 10–100 mg/L = harmful to aquatic organisms; > 100 mg/L = not classified. Data are summarized from the database provided in the Supplemental Information.

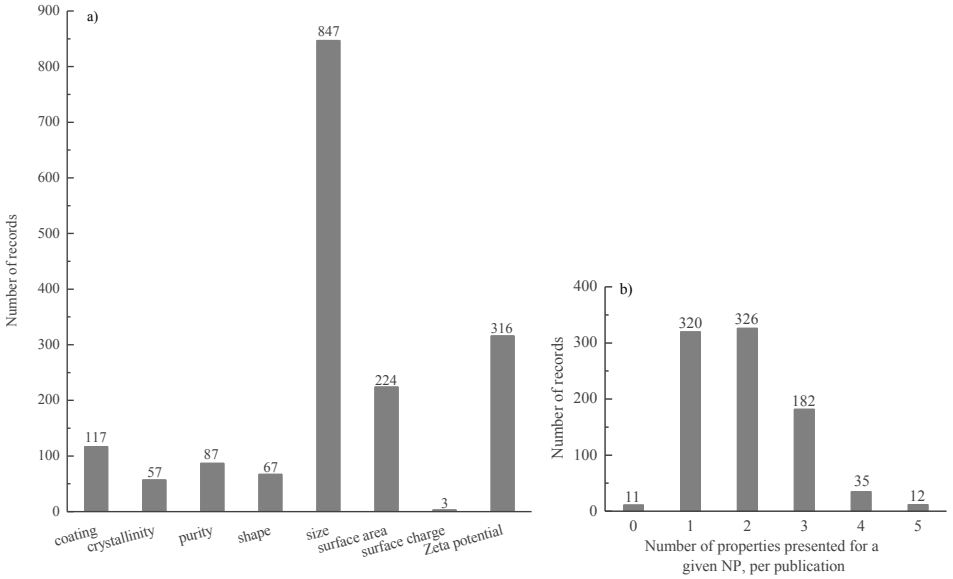


Figure 2.5. NP characterization in the publications retrieved. (a) Shows the number of records with the measured properties in the data assembled; (b) shows the number of properties studied in a given NP, per publication. The data were extracted from the database accessed on 27 February 2014, as shown in the Supplementary Information.

2.4 Outlook

As the number and variety of NPs is expected to increase rapidly, the development of reliable models that allow the prediction of potential toxicity is of vital importance to NP risk assessment. The task of safe-by-design for nanotechnology, amongst others, necessitates the prediction of nanotoxicity based on the pristine structure and basic properties of NPs. The (Q)SAR approach is considered as a possible way forward in this respect. However, several challenges lie ahead regarding a number of vital issues.

2.4.1 Data consistency

Even though 886 toxicity records were retrieved, based on 910 publications from the Web of Science Core Collection, the availability of experimental data on specific toxicity endpoints for nano-(Q)SAR model development remains limited because of poor data

consistency. The data collected is, to some degree, influenced by a range of protocols and experimental conditions, such as the target organisms, type of endpoints, and biological effects. Based on the analysis depicted in Figure 2.3, Table 2.2, and Table S2.5, only growth inhibition (EC50 to *P. subcapitata*) and mortality (LC50 to *D. magna*, *E. coli*, and *D. rerio*) could be potentially modeled with the data retrieved. This finding stems from the fact that only toxicity data generated under consistent experimental conditions for a large variety of NPs are appropriate for (Q)SAR development — e.g. toxicity data generated according to widely accepted and applied guidelines, such as OECD guidelines, US Environmental Protection Agency guidelines. Meanwhile, nano-(Q)SAR modelers could also consider databases like the NBI database when assembling the information of interest.

2.4.2 Data evaluation

Poor quality or unreliable data may lead to models with limited statistical significance or predictivity. Notwithstanding the limitation of data availability, it is to be noted that the use of suitable protocols for evaluating the quality of the toxicity data tested/measured by different methods, and in various laboratories, remains crucial. Previously, different schemes have been described for assessing data quality, and are expected to be interpreted in the light of the purpose for which the data are to be used (Tielemans et al., 2002; Hobbs et al., 2005; Schneider et al., 2009; Klimisch et al., 1997; Przybylak et al., 2012). Specifically, Lubinski et al. (2013) proposed a data quality evaluation framework, with a focus on data applicability to (Q)SARs. These studies offer possible ways of filtering assembled data for the development of nano-(Q)SARs. It is worth noting that a suitable protocol for this task ought to reach a balance between data quality and data availability, ensuring that sufficient data but of good quality could be put into use. In this review, we did not consider the application of a data quality evaluation framework.

2.4.3 Characterization of NPs

NP characterization plays a vital role in the development of nano-(Q)SARs. The obvious first step in nano-(Q)SAR modeling is to link characteristics of pristine NPs to toxicity endpoints. Based on our analysis (Figure 2.5), only a few measured NP properties were provided, and their importance with regard to proper characterisation of pristine NPs is remarkably limited. If more properties of existing NPs could be derived, then there would be a greater possibility that metal-based NPs could be adequately characterized for nano-(Q)SAR modeling. However, challenges to the derivation of adequate descriptors for nano-(Q)SARs still remain, mainly in two areas. First, rather than being characterized as a defined entity, NPs can generally only be defined in a somewhat arbitrary way before being described in terms of descriptors suitable for modeling. Often, NPs are complicated

assemblies, probably coated or functionalized with diverse molecules, the composition of which may vary over time. This makes it impossible to define them strictly as an entity that is interacting with a biological species and causing toxicity. Secondly, the high complexity of the 3-D structure of NPs hinders the calculation of descriptors based on current computational approaches. Uncertainty surrounding the 3-D structure of an NP still exists, even when NP compositions are apparently properly reported (Fourches et al., 2011). These issues pose a big challenge in the feasibility and efficiency of descriptor derivation for nano-(Q)SARs (Gajewicz et al., 2012). Accordingly, descriptors of pristine NPs that describe the essential structural properties without missing crucial structural information and consuming much time for calculation ought to be developed for modeling.

2.4.4 The dynamics of pristine NPs in exposure media

Even though the possibility of building nano-(Q)SAR models based on the characteristics of both pristine and medium-related NPs has already been shown, the feasibility of applying (Q)SARs in nanotoxicity prediction is still largely unknown. According to the (Q)SAR paradigm, it is possible to predict the toxicological effects directly from the physical-chemical properties of the entities of interest (Winkler et al., 2013), which leads to the potential use of (Q)SARs as possible alternative *in silico* screening protocols for testing, without obtaining experiment-related information. However, when in contact with artificial and natural aqueous media, very often the metal-based NPs interact strongly with constituents in the medium (Tiede et al., 2009) and undergo dramatic changes to their surface properties (El Badawy et al., 2010), as well as to their dissolution and aggregation behavior (Baalousha et al., 2008; Tso et al., 2010). These changes affect NP mobility, bioavailability, and ultimately toxicity to organisms (El Badawy et al., 2011; Handy et al., 2008; Hua et al., 2014; Suresh et al., 2013). It should be acknowledged that these interactions are dynamic in nature, and often kinetically rather than thermodynamically controlled, as is usually the case for non-particulate chemicals. Therefore, it is possible that, in some cases, toxicity information can be poorly modeled if the information available is solely based on the characteristics of pristine NPs. Relationships developed between toxicity endpoints and characteristics of pristine NPs without considering the dynamic transformations of NPs in the media, will most likely result in models of low statistical significance, predictability and relevance. In such a context, better interpretation of the dynamic processes influencing NPs in aqueous environments is highly required for modeling and predicting the biological effects of NPs.

2.5 Conclusions

This study identified and collated nanotoxicity data on metal-based NPs, based on the characteristics of pristine NPs. The resulting database, put together from information available in peer-review journals and which will be available as supplementary information on the ATLA website, provides a list of toxicity data of metal-based NPs and should assist toxicologists who work with metal-based NPs.

Our results show that the existing data cannot currently be used to the extent that would be needed to efficiently develop predictive toxicity models for metal-based NPs. Data consistency is shown to play a vital role when performing in-depth quantitative analysis of the experimental data, and numerous data gaps were identified when comparing species and NPs tested. To this end, we recommend that further testing is performed on additional key species and NPs, in order to accurately assess the impacts of metal-based NPs on ecosystems and to develop widely applicable nano-(Q)SARs. It should be emphasized that, to obtain data that will be acceptable for use in further modeling applications, experiments need to be based on consistent experimental conditions or on generally accepted and widely applied guidelines (e.g. OECD or US Environmental Protection Agency guidelines).

We conclude from this review that (Q)SAR approaches have limited potential when used for predicting NP toxicity based on the characteristics of pristine NPs. However, the review nevertheless provides insight into a number of issues vital to the development of nano-(Q)SARs.

Acknowledgements

Guangchao Chen greatly thanks the funding support by the Chinese Scholarship Council (201306060076). Martina G. Vijver is funded by the VIDI-project number 016.141.315 of NWO. Part of the work was performed within the framework of the RIVM sponsored project "IRAN" and the NATO sponsored project "Ecotoxicity of Metal and Metal Oxide Nanoparticles: Experimental Study and Modeling" (project number SFPP·984401).

References

Altschuh J, Bruggemann R, Santl H, Eichinger G, Piringer OG. Henry's law constants for a diverse set of organic chemicals: Experimental determination and comparison of estimation methods. Chemosphere. 1999, 39:1871–1887.

Arnot JA, Gobas FAPC. A review of bioconcentration factor (BCF) and bioaccumulation factor (BAF) assessments for organic chemicals in aquatic organisms. Environ Rev. 2006, 14:257–97.

Baalousha M, Manciulea A, Cumberland S, Kendall K, Lead JR. Aggregation and surface properties of iron oxide nanoparticles: influence of pH and natural organic matter. Environ Toxicol Chem. 2008, 27:1875-82.

Blaise C, Gagné F, Férard JF, Eullaffroy P. Ecotoxicity of selected nano-materials to aquatic organisms. Environ Toxicol. 2008, 23:591-8.

Bondarenko O, Juganson K, Ivask A, Kasemets K, Mortimer M, Kahru A. Toxicity of Ag, CuO and ZnO nanoparticles to selected environmentally relevant test organisms and mammalian cells *in vitro*: A critical review. Arch Toxicol. 2013, 87:1181-200.

Chau YT, Yap CW. Quantitative nanostructure–activity relationship modelling of nanoparticles. RSC Adv. 2012, 2:8489–8496.

Chen G, Li X, Chen J, Zhang YN, Peijnenburg WJ. Comparative study of biodegradability prediction of chemicals using decision trees, functional trees, and logistic regression. Environ Toxicol Chem. 2014, 33:2688-93.

Commission of the European Communities (CEC). *Technical Guidance Document in Support of Commission* Directive 93/67/EEC *on Risk Assessment for New Notified Substances. Part II, Environmental Risk Assessment*. Luxembourg, Luxembourg: Office for Official Publications of the European Communities, 1996.

Cumming JG, Davis AM, Muresan S, Haeberlein M, Chen H. Chemical predictive modelling to improve compound quality. Nat Rev Drug Discov. 2013, 12:948-62.

Ehret J, Vijver M, Peijnenburg W. The application of QSAR approaches to nanoparticles. Altern Lab Anim. 2014, 42:43-50.

El Badawy AM, Luxton TP, Silva RG, Scheckel KG, Suidan MT, Tolaymat TM. Impact of environmental conditions (pH, ionic strength, and electrolyte type) on the surface charge and aggregation of silver nanoparticles suspensions. Environ Sci Technol. 2010, 44:1260-6.

El Badawy AM, Silva RG, Morris B, Scheckel KG, Suidan MT, Tolaymat TM. Surface charge-dependent toxicity of silver nanoparticles. Environ Sci Technol. 2011, 45:283-7.

Epa VC, Burden FR, Tassa C, Weissleder R, Shaw S, Winkler DA. Modeling biological activities of nanoparticles. Nano Lett. 2012, 12:5808-12.

Fourches D, Pu D, Tassa C, Weissleder R, Shaw SY, Mumper RJ, Tropsha A. Quantitative nanostructure–activity relationship modeling. ACS Nano. 2010, 4:5703-12.

Fourches D, Pu D, Tropsha A. Exploring quantitative nanostructure–activity relationships (QNAR) modeling as a tool for predicting biological effects of manufactured nanoparticles. Comb Chem High Throughput Screen. 2011, 14:217-25.

Gajewicz A, Rasulev B, Dinadayalane TC, Urbaszek P, Puzyn T, Leszczynska D, Leszczynski J. Advancing risk assessment of engineered nanomaterials: Application of computational approaches. Adv Drug Deliv Rev. 2012, 64:1663-93.

Gajewicz A, Schaeublin N, Rasulev B, Hussain S, Leszczynska D, Puzyn T, Leszczynski J. Towards understanding mechanisms governing cytotoxicity of metal oxides nanoparticles: Hints from nano-QSAR studies. Nanotoxicology. 2015, 9:313-25.

Ghorbanzadeh M, Fatemi MH, Karimpour M. Modeling the cellular uptake of magnetofluorescent nanoparticles in pancreatic cancer cells: A quantitative structure activity relationship study. Ind Eng Chem Res. 2012, 51:10712–10718.

Gladis F, Eggert A, Karsten U, Schumann R. Prevention of biofilm growth on man-made surfaces: Evaluation of antialgal activity of two biocides and photocatalytic nanoparticles. Biofouling. 2010, 26:89-101.

Handy RD, von der Kammer F, Lead JR, Hassellöv M, Owen R, Crane M. The ecotoxicology and chemistry of manufactured nanoparticles. Ecotoxicology. 2008, 17:287-314.

Hobbs DA, Warne MS, Markich SJ. Evaluation of criteria used to assess the quality of aquatic toxicity data. Integr Environ Assess Manag. 2005, 1:174-80.

Hu X, Cook S, Wang P, Hwang HM. *In vitro* evaluation of cytotoxicity of engineered metal oxide nanoparticles. Sci Total Environ. 2009, 407:3070-2.

Hua J, Vijver MG, Richardson MK, Ahmad F, Peijnenburg WJ. Particle-specific toxic effects of differently shaped zinc oxide nanoparticles to zebrafish embryos (*Danio rerio*). Environ Toxicol Chem. 2014, 33:2859-68.

Ivask A, Juganson K, Bondarenko O, Mortimer M, Aruoja V, Kasemets K, Blinova I, Heinlaan M, Slaveykova V, Kahru A. Mechanisms of toxic action of Ag, ZnO and CuO nanoparticles to selected ecotoxicological test organisms and mammalian cells *in vitro*: A comparative review. Nanotoxicology. 2014, 8:57-71.

Jo HJ, Choi JW, Lee SH, Hong SW. Acute toxicity of Ag and CuO nanoparticle suspensions against *Daphnia magna*: The importance of their dissolved fraction varying with preparation methods. J Hazard Mater. 2012, 227-228:301-8.

Kahru A, Dubourguier HC. From ecotoxicology to nanoecotoxicology. Toxicology. 2010, 269:105-19.

Kahru A, Ivask A. Mapping the dawn of nanoecotoxicological research. Acc Chem Res. 2013, 46:823-33.

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 In Vitro. 2014, 28:600-6.

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. 2014, 107:162-9.

Kleandrova VV, Luan F, González-Díaz H, Ruso JM, Melo A, Speck-Planche A, Cordeiro MN. Computational ecotoxicology: Simultaneous prediction of ecotoxic effects of nanoparticles under different experimental conditions. Environ Int. 2014, 73:288-94.

Klimisch HJ, Andreae M, Tillmann U. A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. Regul Toxicol Pharmacol. 1997, 25:1-5.

Liu R, Rallo R, George S, Ji Z, Nair S, Nel AE, Cohen Y. Classification NanoSAR development for cytotoxicity of metal oxide nanoparticles. Small. 2011, **7**:1118–1126.

Liu R, Zhang HY, Ji ZX, Rallo R, Xia T, Chang CH, Nel A, Cohen Y. Development of structure–activity relationship for metal oxide nanoparticles. Nanoscale. 2013, 5:5644-53.

Luan F, Kleandrova VV, González-Díaz H, Ruso JM, Melo A, Speck-Planche A, Cordeiro MN. Computer-aided nanotoxicology: assessing cytotoxicity of nanoparticles under diverse experimental conditions by using a novel QSTR-perturbation approach. Nanoscale. 2014, 6:10623–10630.

Lubinski L, Urbaszek P, Gajewicz A, Cronin MT, Enoch SJ, Madden JC, Leszczynska D, Leszczynski J, Puzyn T. Evaluation criteria for the quality of published experimental data on nanomaterials and their usefulness for QSAR modelling. SAR QSAR Environ Res. 2013, 24:995-1008.

Maynard AD, Aitken RJ, Butz T, Colvin V, Donaldson K, Oberdörster G, Philbert MA, Ryan J, Seaton A, Stone V, Tinkle SS, Tran L, Walker NJ, Warheit DB. Safe handling of nanotechnology. Nature. 2006, 444:267-9.

Mustad AP, Smeets B, Jeliazkova N, Jeliazkov V, Willighagen EL. *Summary of the Spring 2014 NSC Database Survey*. 2014. Available at: http://figshare.com/articles/Summary_of_the_Spring_2014_NSC_Database_Survey/1195888 (Accessed 16.07.14).

OECD (undated). *Quantitative Structure–Activity Relationships Project [(Q)SARs]*. Paris. France: Organisation for Economic Co-operation and Development. Available at: http://www.oecd.org/env/hazard/qsar (Accessed 16.07.15).

Okse C, Ma CY, Wang XZ. Current situation on the availability of nanostructure-biological activity data. SAR QSAR Environ Res. 2015, 26:79-94.

Pathakoti K, Huang MJ, Watts JD, He X, Hwang HM. 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. 2014, 130:234-40.

Pavan M, Worth AP. *Review of QSAR Models for Ready Biodegradation*, 78pp. Ispra, Italy: European Commission, Joint Research Centre, 2006.

Pavan M, Worth AP, Netzeva TI. *Review of QSAR Models for Bioconcentration*, 125pp. Ispra, Italy: European Commission, Joint Research Centre, 2006.

Przybylak KR, Madden JC, Cronin MT, Hewitt M. Assessing toxicological data quality: basic principles, existing schemes and current limitations. SAR QSAR Environ Res. 2012, 23:435-59.

Puzyn T, Rasulev B, Gajewicz A, Hu X, Dasari TP, Michalkova A, Hwang HM, Toropov A, Leszczynska D, Leszczynski J. Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles. Nat Nanotechnol. 2011, 6:175-8.

Russell WMS, Burch RL. *The Principles of Humane Experimental Technique*, 238pp. London, UK: Methuen, 1959.

Sanderson H, Johnson DJ, Wilson CJ, Brain RA, Solomon KR. Probabilistic hazard assessment of environmentally occurring pharmaceuticals toxicity to fish, daphnids and algae by ECOSAR screening. Toxicol Lett. 2003, 144:383-95.

Schneider K, Schwarz M, Burkholder I, Kopp-Schneider A, Edler L, Kinsner-Ovaskainen A, Hartung T, Hoffmann S. "ToxRTool", a new tool to assess the reliability of toxicological data. Toxicol Lett. 2009, 189:138-44.

Schrand AM, Rahman MF, Hussain SM, Schlager JJ, Smith DA, Syed AF. Metal-based nanoparticles and their toxicity assessment. Wiley Interdiscip Rev Nanomed Nanobiotechnol. 2010, 2:544-68.

Shaw SY, Westly EC, Pittet MJ, Subramanian A, Schreiber SL, Weissleder R. Perturbational profiling of nanomaterial biologic activity. Proc Natl Acad Sci U S A. 2008, 105:7387-92.

Singh KP, Gupta S. Nano-QSAR modeling for predicting biological activity of diverse nanomaterials. RSC Adv. 2014, 4:13215–13230.

Sizochenko N, Rasulev B, Gajewicz A, Kuz'min V, Puzyn T, Leszczynski J. From basic physics to mechanisms of toxicity: The "liquid drop" approach applied to develop predictive classification models for toxicity of metal oxide nanoparticles. Nanoscale. 2014, 6:13986-93.

Song L, Vijver MG, Peijnenburg WJ, De Snoo GR. Smart nanotoxicity testing for biodiversity conservation. Environ Sci Technol. 2011, 45:6229-30.

Suresh AK, Pelletier DA, Doktycz MJ. Relating nanomaterial properties and microbial toxicity. Nanoscale. 2013, 5:463-74.

Tiede K, Hassellöv M, Breitbarth E, Chaudhry Q, Boxall AB. Considerations for environmental fate and ecotoxicity testing to support environmental risk assessments for engineered nanoparticles. J Chromatogr A. 2009, 1216:503-9.

Tielemans E, Marquart H, De Cock J, Groenewold M, Van Hemmen J. A proposal for evaluation of exposure data. Ann Occup Hyg. 2002, 46:287-97.

Toropov AA, Toropova AP, Benfenati E, Gini G, Puzyn T, Leszczynska D, Leszczynski J. Novel application of the CORAL software to model cytotoxicity of metal oxide nanoparticles to bacteria *Escherichia coli*. Chemosphere. 2012, 89:1098-102.

Toropov AA, Toropova AP, Puzyn T, Benfenati E, Gini G, Leszczynska D, Leszczynski J. QSAR as a random event: Modeling of nanoparticles uptake in $PaCa_2$ cancer cells. Chemosphere. 2013, 92:31-7.

Tso CP, Zhung CM, Shih YH, Tseng YM, Wu SC, Doong RA. Stability of metal oxide nanoparticles in aqueous solutions. Water Sci Technol. 2010, 61:127-33.

Winkler DA, Mombelli E, Pietroiusti A, Tran L, Worth A, Fadeel B, McCall MJ. Applying quantitative structure-activity relationship approaches to nanotoxicology: Current status and future potential. Toxicology. 2013, 313:15-23.

Weissleder R, Kelly K, Sun EY, Shtatland T, Josephson L. Cell-specific targeting of nanoparticles by multivalent attachment of small molecules. Nat Biotechnol. 2005, 23:1418-23.

Zhang H, Ji Z, Xia T, Meng H, Low-Kam C, Liu R, Pokhrel S, Lin S, Wang X, Liao YP, Wang M, Li L, Rallo R, Damoiseaux R, Telesca D, Mädler L, Cohen Y, Zink JI, Nel AE. Use of metal oxide nanoparticle band gap to develop a predictive paradigm for oxidative stress and acute pulmonary inflammation. ACS Nano. 2012, 6:4349-68.

Chapter 2 Supplemental Information

Table S2.1. The key words used to select the test species for the data search

Tested species	Key words
bacteria	bacter* OR Escherichia* OR Staphylococcus* OR Bacillus*
yeast	yeast* OR Candida* OR fungi* OR Saccharomyces*
algae	*alga* OR Pseudokirchneriella* OR Chlamydomonas*
protozoa	protozoa* OR Paramecium* OR Tetrahymena*
crustacean	crustacea* OR daphni* OR Thamnocephalus*
nematode	nematode* OR Caenorhabditis*
fish	*fish* OR Oryzias* OR Pimephales* OR Danio*

In the search query of each group of organism, a general key word characterizing the organism was firstly considered (i.e. bacter*, yeast*, *alga*, protozoa*, crustacea*, nematode*, and *fish*). A further search with other key words referring to different species was subsequently carried out to enclose some studies which addressed the nanotoxicity to these organisms but did not use the key words bacter*, yeast*, *alga*, protozoa*, crustacea*, nematode* or *fish* in either the title, abstract or key words of the publications. The extra key words were chosen empirically.

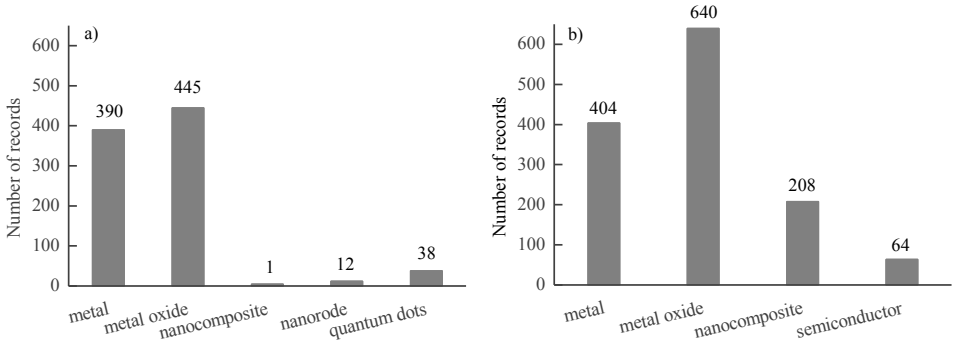


Figure S2.1. Comparison of types of NPs in our database (left) and the NBI database (right). (a) Distribution of NP types in our database; (b) Distribution of NP types in the NBI database. Number of records are shown in the figures.

Table S2.2. Key words characterizing the metal-based nanoparticles (NPs) for data search

NPs	Key words
Ag	nano* AND Silver* OR "nano* AND Ag*"
Al	nano* AND Aluminum* OR Al2O3 OR "nano* AND Al*"
Au	nano* AND gold* OR "nano* AND Au*"
Bi	nano* AND Bismuth*
Cd	nano* AND Cadmium* OR CdO OR "nano* AND Cd*"
Ce	nano* AND Cerium* OR CeO2
Co	nano* AND Cobalt* OR Co3O4 OR "nano* AND Co*"
Cr	nano* AND Chromium* OR CrO3 OR "nano* AND Cr*"
Cu	nano* AND Copper* OR CuO OR "nano* AND Cu*"
Fe	nano* AND Iron* OR Fe2O3 OR Fe3O4 OR "nano* AND Fe*"
In	nano* AND Indium* OR In2O3
La	nano* AND Lanthanum* OR La2O3
Mn	nano* AND Manganese* OR MnO OR Mn3O4 OR "nano AND Mn*"
Mo	nano* AND Molybdenum* OR MoO3 OR "nano AND Mo*"
Ni	nano* AND Nickel* OR NiO OR "nano* AND Ni*"
Pt	nano* AND Platinum* OR PtO2 OR "nano* AND Pt*"
Sb	nano* AND Antimony* OR Sb2O3 OR "nano* AND Sb*"
Se	nano* AND Selenium* OR SeO2 OR "nano* AND Se*"
Si	nano* AND Silic* OR SiO2
Ti	nano* AND Titanium* OR TiO2 "nano* AND Ti*"
V	nano* AND Vanadium* OR V2O5
Zn	nano* AND Zinc* OR ZnO "nano* AND Zn*"
Zr	nano* AND Zirconium* OR ZrO2 OR "nano* AND Zr*"

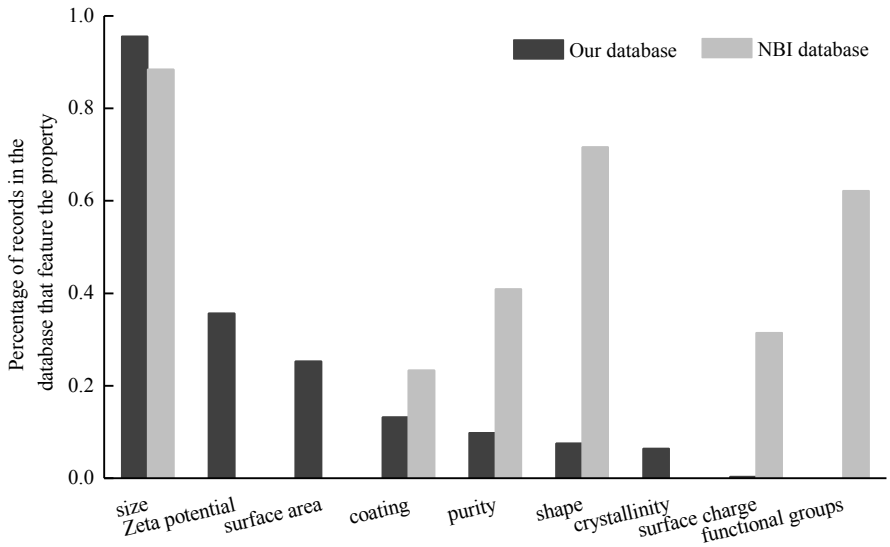


Figure S2.2. Comparison of characterization of NPs in respective database.

Table S2.3. The numbers of papers retrieved from the Web of Science Core Collection by condition (a). The search was carried out on 27 February 2014

		Bacteria	**Yeast**	**Algae**	**Protozoa**	**Crustacean**	**Nematode**	**Fish**	**Total**
1	Ag	219	29	40	7	63	27	112	**394**
2	Al	16	6	9	2	10	5	9	**39**
3	Au	41	21	10	1	14	3	37	**110**
4	Bi	0	0	0	0	0	0	2	**2**
5	Cd	18	11	9	2	18	5	25	**74**
6	Ce	26	0	6	1	11	3	11	**41**
7	Co	9	3	3	0	1	0	5	**18**
8	Cr	6	2	0	0	1	0	1	**10**
9	Cu	62	14	35	4	42	5	45	**148**
10	Fe	41	13	9	4	9	1	13	**76**
11	In	2	0	1	0	2	0	1	**3**
12	La	3	0	0	0	0	1	0	**3**
13	Mn	6	1	1	0	2	0	1	**8**
14	Mo	2	0	0	0	0	0	0	**2**
15	Ni	12	2	4	0	2	0	5	**20**
16	Pt	3	3	0	0	0	0	1	**7**
17	Sb	3	1	0	0	0	0	0	**3**
18	Se	6	3	1	0	4	0	7	**16**
19	Si	41	14	11	1	14	2	21	**90**
20	Ti	82	13	42	6	99	10	88	**249**
21	V	1	0	0	0	1	0	1	**2**
22	Zn	67	13	39	3	42	9	25	**143**
23	Zr	3	0	0	0	0	0	1	**4**
	Total	**445**	**114**	**141**	**19**	**200**	**46**	**259**	**982**

Table S2.4. The numbers of papers focusing on different metal-based NPs and retrieved with condition (b)

NPs	**Ag**	**Al**	**Au**	**Bi**	**Cd**	**Ce**	**Co**	**Cr**	**Cu**	**Fe**	**In**	**La**
Number of papers	383	35	107	2	72	40	18	5	139	70	3	2

NPs	**Mn**	**Mo**	**Ni**	**Pt**	**Sb**	**Se**	**Si**	**Ti**	**V**	**Zn**	**Zr**
Number of papers	2	2	14	6	1	16	79	238	1	137	3

Table S2.5. An overview of the diversity analysis of the metal-based NPs corresponding to relevant toxicity endpoints and tested organisms, as summarized from the database

		Danio rerio	**Vibrio fischeri**	**Staphylococcus aureus**	**Pseudokirchneriella**	**Escherichia coli**	**Daphnia magna**
Mortality	LC50	Ag (5); CdSe (1); Cu (3); Cu_2O (1);				Co_3O_4 (2); CuO (3); TiO_2 (10);	Ag (25); Al_2O_3 (1); CeO_2 (4);
	LOEC						Ag (1); TiO_2 (4); CeO_2 (3)
	NOEC						Ag (1); CeO_2 (3); TiO_2 (4); ZnO
	MBC	Ag (4)		Ag (15); Cu (3)		Ag (16); Cu (4)	
	Other	Ag (2)					CuO (1); TiO_2 (4); ZnO (3)
Growth inhibition	EC10				Au (5); CeO_2 (6); SiO_2 (2); TiO_2		
	EC20				Au (2); CeO_2 (6); SiO_2 (2); TiO_2		
	EC50				Al (1); Ag (11); Au (5); CeO_2	Ag (1); CdS (1); CuO (1); NiO	
	LOEC				CeO_2 (3); SiO_2 (2)	CuO (2); NiO (1); ZnO (1)	
	MIC			Ag (27); Cu (3); ZnO (1)	CeO_2 (3); SiO_2 (2); CuO (1);	Fe (5); Ag (20); Cu (8); Fe_3O_4	
	NOEC					NiO (1); Sb_2O_3 (1)	
Immobilization	EC10						Ag (6)
	EC50						Ag (22); Al_2O_3 (1); CuO (13);
	NOEC						CuO (1); ZnO (1); TiO_2 (1)

Table S2.5. (Continued)

		Danio rerio	Vibrio fischeri	Staphylococcus aureus	Pseudokirchneriella	Escherichia coli	Daphnia magna
Luminescence inhibition	EC10		Ag (2)				
	EC20		Au (6); CdSe/ZnS (5);				
	EC50		Ag (4); Au (6); CdSe/ZnS (5);			Ag (14); CuO (12); ZnO (12)	
	MIC		CuO (1); ZnO (1)				
	LOEC		TiO_2 (2)				
	NOEC		CuO (1); ZnO (1)				
Malformation	EC50	Ag (10)				Ag (1); TiO_2 (1); ZnO (3)	
	NOEC					ZnO (2)	
Reproduction inhibition	EC10						CeO_2 (3); TiO_2 (1); ZnO (1)
	EC20						CeO_2 (3); ZnO (1)
	EC50				TiO_2 (2)		CeO_2 (3); TiO_2 (2); ZnO (3)
	LOEC						CeO_2 (3); TiO_2 (1)
	NOEC						CeO_2 (3); TiO_2 (1)
Others		Ag (2); ZnO (1); Fe_2O_3 (2)					ZnO (3)

D. magna - Daphnia magna; *E. coli - Escherichia coli*; *P. subcapitana - Pseudokirchneriella subcapitata*; *S. aureus - Staphylococcus aureus*; *V. fischeri - Vibrio fischeri*; *D. rerio = Danio rerio.* The record number is given in parentheses.

Table S2.6. Median L(E)C50 values of metal-based NPs to tested species. The number of toxicity records is indicated in parentheses

	Algae	**Bacteria**	**Crustacean**	**Fish**	**Nematode**	**Protozoa**	**Yeast**
Ag	0.10 (25)	19.25 (20)	0.01 (57)	1.36 (33)	2.85 (19)	38.00 (5)	
Al	20.86 (3)	172.83 (1)	45.79 (7)		43.25 (1)		
Au	38.00 (5)	0.32 (6)					38.0 (1)
Ce	7.07 (18)	28.96 (12)	29.89 (7)				
Co		33.08 (2)	1.67 (1)				
Cu	2.80 (4)	17.36 (20)	0.61 (22)	0.83 (4)		127.00 (3)	4.38 (13)
Fe		309.81 (2)	36.00 (3)	37.35 (3)		0.57 (1)	
La		388.37 (1)					
Ni	25.50 (3)	111.43 (2)	2.28 (2)				
Sb		344.57 (2)					
Se				1.00 (1)			
Si				112.80 (1)		208.02 (1)	
Sn		826.02 (1)					
Ti	18.96 (27)	111.00 (16)	3.90 (31)	93.00 (13)	47.94 (1)		
Zn	1.94 (2)	24.80 (27)	1.20 (15)	1.543 (4)	635.00 (5)	5.00 (3)	100.80 (2)

CHAPTER 3

RECENT ADVANCES TOWARDS THE DEVELOPMENT OF (QUANTITATIVE) STRUCTURE-ACTIVITY RELATIONSHIPS FOR METALLIC NANOMATERIALS: A CRITICAL REVIEW

Chen G, Vijver MG, Xiao Y, Peijnenburg WJGM

Submitted to *Materials*. Under revision

Abstract

The exponential increase of nanotechnology has raised concerns on the risks posed by engineered nanomaterials (ENMs). Recent studies on the ecotoxicity of ENMs addressed that these materials could potentially cause adverse effects to human health and to biota. A comprehensive assessment of ENMs' risks is thus urgently needed, which is, however, severely hindered by time, financial burden, and ethical considerations. Gathering the required information in a fast and inexpensive way seems essential. In such a context, the extension of the conventional (quantitative) structure-activity relationships ((Q)SARs) approach to nanotoxicology, i.e. nano-(Q)SARs, is a possible solution. Recently, various attempts have been made to correlate ENMs' characteristics to the biological effects elicited by ENMs. This highlighted the potential applicability of (Q)SAR in the nanotoxicity field to aid in prioritizing information on nanotoxicity and in rationalizing the risk assessment of ENMs. This review summarizes and discusses the current knowledge on nano-(Q)SARs for metallic ENMs with regard to the aspects (i) sources of data; (ii) existing nano-(Q)SARs; (iii) mechanistic interpretation; and (iv) an outlook on the further development of this frontier. The review aims to present key advances in relevant nano-modeling studies and to stimulate future research efforts in this quickly developing field of research.

Key words: cellular uptake, metallic, nanomaterials, (Q)SARs, toxicity

3.1 Introduction

Manipulating matter at the nanoscale (1-100 nm) has provided a way forward to designing materials that exhibit inimitable magnetic, electrical, optical, and thermal properties compared to the bulk counterparts (Puzyn et al., 2009). The products of engineered nanomaterials (ENMs) are consequently finding routine use in a wide range of commercial applications (Linkov et al., 2009). It was expected that the exponentially growing nano-market would reach a turnover of $65 billion by 2019 (Winkler, 2016). The release of ENMs into landfills, air, surface waters, and other environmental compartments therefore seems inevitable. In such a context, it is very likely for humans and for biota to encounter these nano-products and to be at risk given the potential adverse effects induced by ENMs. Studies on the cytotoxicity (Asare et al., 2012; Nirmala et al., 2011; Wiesner et al., 2006), neurotoxicity (Long et al., 2006; Win-Shwe and Fujimaki, 2011; Wu et al., 2011), genotoxicity (Asare et al., 2012; Kumari et al., 2011; Sharma et al., 2011), and ecotoxicity (Ellegaard-Jensen et al., 2012; Thill et al., 2006; Tran et al., 2010) of ENMs have shown that, miniaturization of materials to the nanoscale may result in the appearance of evident ENM toxicity on organisms and human cell lines, which does not always occur at the bulk scales. This highlighted the potential risks associated with the fast developing field of nanotechnology. Hence, seeking ways for the risk assessment of ENMs becomes imperative.

According to the commonly accepted procedures of chemical risk assessment, both exposure and hazard assessment are key to evaluate the risks of ENMs (Gajewicz et al., 2012; Worth, 2010). Hazard characterization, which aims at defining the dose-responses for targets or target-species is supposed to be mainly derived according to standardized test guidelines (e.g., Organization for Economic Co-operation and Development (OECD) guidelines). However, despite the existence of these powerful testing protocols, the possibility of covering all the existing and also newly synthesized ENMs in the "nano pool" is reduced taking into account the need of cost-effectiveness testing whilst minimizing the use of test animals. Considering the exponential increase of nanotechnology, the scarcity of data on ENM toxicity poses a major barrier to perform comprehensive hazard assessment of ENMs. As a result, development of fast and inexpensive alternative approaches filling the data gaps and assisting in rationalizing ENMs' risk assessment is of significant importance. Moreover, the principle of the 3R (replacement, reduction, and refinement) rule also calls for a reduction in the animal use and developing alternative non-animal testing approaches (Puzyn et al., 2011; Russell and Burch, 1959).

One of the most promising approaches that has long been particularly helpful for predicting biological effects of chemicals is the (quantitative) structure-activity relationship ((Q)SAR) method (Fernández et al., 2012; Kar and Roy, 2010; 2012). The (Q)SAR approach enables

the encoding of existing knowledge into predictive models which directly correlate the molecular structure with toxicity of a chemical. The role of (Q)SARs in predictive toxicology is (Peijnenburg, 2009; Raymond et al., 2001):

(i) To provide fast and inexpensive high-throughput screening methods estimating the toxicity of chemical entity;

(ii) To assist the classification of chemicals according to their toxicity;

(iii) To help understand the underlying toxic mechanisms.

Two issues especially figure in the extraction of meaningful relationships between structures and biological effects to yield (Q)SAR models: the so-called molecular descriptor (measured or calculated) characterizing vital structural information of chemicals, and the so-called endpoint describing the biological effects of interest (Ivask et al., 2014). According to the OECD Principles for (Q)SAR Validation (OECD, 2007), it is essential for a (Q)SAR model considered suited for regulatory purposes to include information on: (i) a defined endpoint; (ii) an unambiguous algorithm; (iii) a defined domain of applicability; (iv) appropriate measures of goodness-of-fit, robustness, and predictivity; and (v) a mechanistic interpretation, if possible.

Facing the strong need of extending the conventional (Q)SAR approach to nanotoxicology, some attempts have been made to link ENMs' biological effects with the characteristics of ENMs. A summary of recent advances towards this field is thus presented in Table 3.1 to offer an overview of the research achievements obtained so far. The underlying literature search was performed by means of an Advanced Search in the Web of Science™ Core Collection on the 22th of February, 2017. The search was manually supplemented with relevant publications not included in the search records. The query is ((((TS=(nano* AND metal)) AND (TS=(toxic*))) AND (TS=(quantitative *structure activity relationship) OR TS=(*QSAR) OR TS=(QNAR) OR TS=(predict*) OR TS=(computation*) OR TS=(model*)))), where the field tag TS refers to the topic of a publication. As can be seen, various nano-(Q)SARs were constructed based on a variety of modeling techniques such as linear and nonlinear regression, support vector machine (SVM), artificial neural networks (ANN), and *k* nearest neighbor (*k*NN). Distinct biological responses induced by ENMs such as cellular uptake and cytotoxicity in different cell lines, and the ecotoxicity of ENMs were addressed. The studies provided in Table 3.1 highlight the potential of (Q)SAR methods to be adopted as a tool in predicting nanotoxicity. Thus, to provide an overview of recent key advances in this field, the state-of-the-art of reported nano-(Q)SARs is discussed on the following aspects: (i) sources of data for modeling; (ii) existing nano-(Q)SARs; (iii)

mechanistic interpretation; and (iv) an outlook on the further development of nano-(Q)SARs identifying major research gaps in the field.

Table 3.1. Overview of the peer-reviewed literatures on nano-(Q)SARs, as generated by means of an advanced literature search in the Web of Science™ Core Collection on 22th of February, 2017, and supplemented with a manual collection of relevant publications not included in the search record. Apart from the references obtained, a general description is given for the models reported

Reference	Description
Burello and Worth, 2011	A model was proposed to show that the oxidative stress potential of metal oxide ENMs could be possibly predicted by looking at the their band gap energy
Chau and Yap, 2012	Developed a final consensus model based on top 5 candidate models constructed by naive Bayes, logistic regression, *k*-nearest neighbor (*k*NN), and support vector machine (SVM), predicting the cellular uptake of 105 ENMs (single metal core) by PaCa2 pancreatic cancer cells
Chen et al., 2016	Global classification models were developed to predict the ecotoxicity of metallic ENMs to different species; classification models were also built for *Danio rerio*, *Daphnia magna*, *Pseudokirchneriella subcapitata*, and *Staphylococcus aureus*
Epa et al., 2012	Modeled (i) cytotoxicity of 31 ENMs to vascular smooth muscle cells based on multiple linear regression and Bayesian regularized artificial neural network; (ii) cellar uptake of 108 ENMs in human umbilical vein endothelial cells (HUVEC) and PaCa2 cells using multiple linear regression with expectation maximization method
Fourches et al., 2010	Generated models predicting (i) cytotoxicity of 44 ENMs with diverse metal cores using SVM method; (ii) cellular uptake of 109 ENMs in PaCa2 cells using *k*NN method
Gajewicz et al., 2015	Applied the multiple linear regression method combined with a genetic algorithm to describe the toxicity of 18 metal oxide ENMs to the human keratinocyte cell line (HaCaT)
Ghorbanzadeh et al., 2012	Cellular uptake of 109 magnetofluorescent ENMs in PaCa2 cells was modeled using multiple linear regression and multilayered perceptron neural network, descriptor selection was performed by combining the self-organizing map and stepwise multiple linear regression
Kar et al., 2014a	Developed a model establishing the cellular uptakes of 109 magnetofluorescent ENMs in PaCa2 cells
Kar et al., 2014b	Using the toxicity dataset of 17 metal oxide ENMs to *Escherichia coli* (*E. coli*), models were built with the multiple linear regression and partial least squares methods
Kleandrova et al., 2014	Perturbation model was introduced for the prediction of ecotoxicity and cytotoxicity of ENMs; molar volume, electronegativity, polarizability, size of the particles, hydrophobicity, and polar surface area were involved in the model
Liu et al., 2011	Classification-models (logistic regression) were developed to predict the cytotoxicity of nine ENMs to the transformed bronchial epithelial cells (BEAS-2B)

Table 3.1. (Continued)

Liu et al., 2013a	A nano-SAR was developed classifying 44 iron-based ENMs into bioactive or inactive, using a naive Bayesian classifier based on 4 descriptors: primary size, spin-lattice and spin-spin relaxivities, and zeta potentials
Liu et al., 2013b	SVM nano-SAR model was constructed on basis of the cytotoxicity data of 24 metal oxide ENMs to BEAS-2B cells and murine myeloid (RAW 264.7) cells
Luan et al., 2014	Perturbation model was presented predicting the cytotoxicity of ENMs against several mammalian cell lines; influence of molar volume, polarizability, and size of the particles were indicated
Mu et al., 2016	A quantitative model was developed based on the toxicity data of 16 metal oxide ENMs to *E. coli* using enthalpy of formation of a gaseous cation (ΔH_{Me+}) and polarization force (Z/r). The toxicity of 35 other metal oxide ENMs was predicted and depicted in the periodic table
Pan et al., 2016	Models were constructed to predict (i) the toxicity of 17 metal oxide ENMs to *E. coli*; (ii) cytotoxicity in HaCaT cells of 18 different metal oxide ENMs. The factors of molecular weight, cationic charge, mass percentage of metal elements, individual and aggregation sizes were discussed
Papa et al., 2015	Cytotoxicity of TiO_2 and ZnO ENMs were modeled by LMR and C4.5 algorithm
Pathakoti et al., 2014	Toxicity and photo-induced toxicity of 17 metal oxide ENMs to *E. coli* was assessed using a self-written least-squares fitting program
Puzyn et al., 2011	Predicted cytotoxicity of 17 metal oxide ENMs to *E. coli* with only one descriptor: enthalpy of formation of a gaseous cation having the same oxidation state as that in the metal oxide structure
Singh and Gupta, 2014	Predictive models were built based on (i) cytotoxicity of different ENMs (with diverse metal cores) in four cell lines (endothelial and smooth muscle cells, monocytes, and hepatocytes); (ii) cellular uptake of 109 ENMs in PaCa2 cells; (iii) cytotoxicity of 17 different metal oxide ENMs to *E. coli*
Sizochenko et al., 2014	Based on random forest regression, developed predictive classification models for (i) toxicity of 17 metal oxide ENMs to *E. coli*; (ii) cytotoxicity of 18 metal oxide ENMs to HaCaT cells
Sizochenko et al., 2015	Structure-activity relationship models (random forest) were introduced for toxicity of 24 metal oxide ENMs towards BEAS-2B and RAW 264.7 cell lines
Toropov et al., 2012	Estimated toxicity of 17 metal oxide ENMs to *E. coli* by employing the SMILES-based optimal descriptors
Toropov et al., 2013	Cellular uptake of 109 ENMs with the same core but different surface modifiers in the PaCa2 cells was modeled based on SMILES-based optimal descriptors
Zhang et al., 2012	A classification model was built for 24 metal oxide ENMs based on the dissolution of metals and energy of conduction band (E_c)

3.2 Sources of data for modeling

As a data-driven approach, the field of nano-(Q)SARs highly relies on generating or assembling qualified experimental data. To integrate the existing information obtained from the various datasets that were successfully used in nano-QSARs, and therefore to aid further studies of nano-modeling, the underlying experimental data in the nano-(Q)SARs mentioned in Table 3.1 were analyzed. As can be seen in Table 3.2, research attention was found to be mainly on the cellular uptake of ENMs by different cell lines, on cytotoxicity, and on the toxicity of ENMs to *Escherichia coli* (*E. coli*). Despite the numerous nano-related tests that are being carried out, it is to be concluded that only a few datasets (with data variety and consistency) were generally used as the data source for nano-(Q)SARs developed so far. The most widely applied data in QSAR-like studies (Table 3.2) are from Weissleder et al. (2005), Puzyn et al. (2011), and Shaw et al. (2008). These experimental datasets are presented and arranged in the order of cellular uptake, cytotoxicity in cell lines, and toxicity to *E. coli* concerning the following aspects: (when available) types and numbers of ENMs, targets or target-species, toxicity endpoints, characteristics of the ENMs provided, and accessibility of relevant information.

3.2.1 Cellular uptake assays

Weissleder et al. (2005) modified the surface of monocrystalline magnetic ENMs (3-nm core of $(Fe_2O_3)_n(Fe_3O_4)_m$) with 146 various small molecules (modifiers) and created a library of 146 water-soluble, magnetic and fluorescent ENMs. ENMs were made magneto-fluorescent by adding the fluorescein isothiocyanate to the ENM surfaces. Uptake of these ENMs by five cell lines was screened afterwards. The cell lines used include pancreatic cancer cells (PaCa2), a macrophage cell line (U937), resting primary human macrophages, activated primary human macrophages, and human umbilical vein endothelial cells (HUVEC). A diversity of cellular uptake of various functionalized ENMs and a high dependence of ENM uptake on the composition of their surface were observed especially in the PaCa2 cells (Chau and Yap, 2012; Fourches et al., 2011). Data on PaCa2 cellular uptake of ENMs can be retrieved from Fourches' studies (Fourches et al., 2010; 2011) and also the studies of Chau and Yap (2012), Kar et al. (2014a), and Ghorbanzadeh et al. (2012). In the absence of data on calculated descriptors for the whole dataset, methods of characterizing ENMs in previous studies are presented in Table 3.3. An analysis of the methods reported in literature shows that emphasis in ENM characterization was so far largely put on the characteristics of ENM surface modifiers, given the conclusion of Weissleder et al. (2005) that the PaCa2 cellular uptake of ENMs highly depends on the surface modification of the ENMs. Descriptor calculation of the modifiers was performed within different softwares (e.g., PaDEL-Descriptor, DRAGON, ADRIANA) providing various molecular descriptors.

Table 3.2. Summary of the experimental data of ENMs used in nano-(Q)SAR studies

nano-(Q)SAR	Dataset used	Number of ENMs	Core of ENMs	Tested organism
Kar et al., 2014b Mu et al., 2016 Pan et al., 2016 Puzyn et al., 2011 Singh and Gupta, 2014 Sizochenko et al., 2014 Toropov et al., 2012	Puzyn et al., 2011	17	Metal oxide	*E. coli*
Chau and Yap, 2012 Epa et al., 2012 Fourches et al., 2010 Ghorbanzadeh et al., 2012 Kar et al., 2014a Singh and Gupta, 2014 Toropov et al., 2013	Weissleder et al., 2005	146	Metal oxide	PaCa2 pancreatic cancer cells
Epa et al., 2012 Fourches et al., 2010 Liu et al., 2013a Singh and Gupta, 2014	Shaw et al., 2008	50	Metal oxide and quantum dots	Endothelial cells, vascular smooth muscle cells, human HepG2 cells, RAW 264.7 cells
Gajewicz et al., 2015 Pan et al., 2016 Sizochenko et al., 2014	Gajewicz et al., 2015	18	Metal oxide	HaCaT cells
Liu et al., 2013b Sizochenko et al., 2015 Zhang et al., 2012	Zhang et al., 2012	24	Metal oxide	BEAS-2B cells; RAW 264.7 cells
Liu et al., 2011	Liu et al., 2011	9	Metal oxide	BEAS-2B cells
Papa et al., 2015	Sayes and Ivanov, 2010	24 TiO_2, 18 ZnO ENMs	TiO_2, ZnO ENMs	Rat L2 lung epithelial cells; rat lung alveolar macrophages
Pathakoti et al., 2014	Pathakoti et al., 2014	17	Metal oxide	*E. coli*
Burello and Worth, 2011 Chen et al., 2016 Kleandrova et al., 2014 Luan et al., 2014			Others	

Table 3.3. Overview of reported information of the data published by Weissleder et al. (2005)

Reference	Method of ENM characterization	Data accessibility	ENM number	Other information
Weissleder et al., 2005			146	Molecular weight and structures
Chau and Yap, 2012	679 one-dimensional (1D), two-dimensional (2D) chemical descriptors of modifiers were calculated using PaDEL-Descriptor (v2.8)	Values of PaCa2 cellular uptake were available (unit: number of ENMs per cell)	109	SMILES (simplified molecular input line entry system)
Epa et al., 2012	691 molecular descriptors of modifiers from DRAGON (v5.5), ADRIANA (v2.2) and an in-house modeling software package		108	List of modifiers
Fourches et al., 2010	MOE descriptors for modifiers were used, including physical properties, surface areas, atom and bond counts, Kier & Hall connectivity indices, kappa shape indices, adjacency and distance matrix descriptors, pharmacophore feature descriptors, and molecular charges	Values of PaCa2 cellular uptake were available (log_{10}[ENM]/cell pM)	109	SMILES
Ghorbanzadeh et al., 2012	Hyperchem program (v7) for constructing molecular structure of modifiers; geometry was optimized with the Austin Model 1 (AM1) semiempirical method; DRAGON for descriptor calculation	Values of PaCa2 cellular uptake were available (log_{10}[ENM]/cell pM)	109	List of modifiers and SMILES
Kar et al., 2014a	A pool of 307 descriptors of modifiers was calculated using Cerius 2 (v4.10), DRAGON 6 and PaDEL-Descriptor (v2.11)	Values of PaCa2 cellular uptake were available (log_{10}[ENM]/cell pM)	109	List of modifiers
Singh and Gupta, 2014	174 molecular descriptors for the modifiers (topological, electronic, geometrical, and constitutional) were calculated using Chemistry Development Kit (CDK v1.0.3)		109	List of modifiers, chemical structures and SMILES
Toropov et al., 2013	SMILES-based optimal descriptors were used		109	SMILES, correlation weights (CWs) of SMILES attributes (SA)

3.2.2 Toxicity to various cell lines

One of the most widely used cell line-based toxicity data for ENMs is from the work of Shaw et al. (2008). In their study, four cell-based assays were performed based on four cell types at four different doses. The four types of cells namely endothelial cells (human aorta),

vascular smooth muscle cells (human coronary artery), hepatocytes (human HepG2 cells), and murine RAW 264.7 leukemic monocyte/macrophage cells were employed to assess the cytotoxicity of 50 ENMs (iron-based ENMs, pseudocaged ENMs, and quantum dots). The four cell-based assays were mitochondrial membrane potential, adenosine triphosphate (ATP) content, apoptosis and reducing equivalents assays. Concentrations of 0.01, 0.03, 0.1, and 0.3 mg/mL Fe for iron-based ENMs, and 3, 10, 30, or 100 nM for quantum dots were used. The ENMs were characterized by their coating, surface modification, size, the spin-lattice (R1) and spin-spin (R2) relaxivities, and the zeta potential. Experimental values were expressed in units of standard deviations of the distribution assessed when cells were only treated with PBS (*Z* score). Fourches et al. (2010) afterwards transformed the 64 features (4 assays × 4 cell lines × 4 doses) of 48 iron-based ENMs into 1 by calculating their arithmetic mean (Z_{mean}) which enabled binary classification studies based on this dataset (data are accessible in the original paper).

Gajewicz et al. (2015) tested the cytotoxicity of 18 metal oxide ENMs to the human keratinocyte cell line (HaCaT). ENMs covered in the dataset include aluminum oxide (Al_2O_3), bismuth oxide (Bi_2O_3), cobalt oxide (CoO), chromic oxide (Cr_2O_3), ferric oxide (Fe_2O_3), indium oxide (In_2O_3), lanthanum oxide (La_2O_3), manganese oxide (Mn_2O_3), nickel oxide (NiO), antimony oxide (Sb_2O_3), silicon dioxide (SiO_2), tin oxide (SnO_2), titanium oxide (TiO_2), vanadium oxide (V_2O_3), tungsten oxide (WO_3), yttrium oxide (Y_2O_3), zinc oxide (ZnO), and zirconium oxide (ZrO_2) ENMs. The cytotoxicity of these ENMs was characterized by cell viability of HaCaT and was expressed in terms of LC50 (concentration of the ENMs that leads to 50% fatality). Experimental data are accessible in the original publication. Moreover, 18 quantum-mechanical and 11 image descriptors were calculated for modeling purposes (Table 3.4). Information on the (aggregation) size for this dataset was provided by Sizochenko et al. (2014) as shown in Table 3.5. Size (50 nm) and aggregation size (180 nm) of WO_3 are not included in the table due to its absence in other datasets depicted in Table 3.5.

By measuring the plasma-membrane leakage via Propidium Iodide (PI) uptake in transformed bronchial epithelial cells (BEAS-2B), Liu et al. (2011) studied the cytotoxicity of a variety of ENMs: Al_2O_3, cerium oxide (CeO_2), Co_3O_4, TiO_2, ZnO, copper oxide (CuO), SiO_2, Fe_3O_4, and WO_3 ENMs. The cytotoxicity was expressed in terms of percentage of membrane-damaged cells (data available in the supplemental information of the original publication). Descriptors calculated were number of metal and oxygen atoms (N_{Metal} and N_{Oxygen}), atomic mass of the ENM metal (m_{Me}), molecular weight of the metal oxide (m_{MeO}), group and period of the ENM metal (G_{Me} and P_{Me}), atomization energy of the metal oxide (E_{MeO}), ENM primary size (d), zeta potential, and isoelectric point (IEP).

Table 3.4. Overview of quantum-mechanical and image descriptors of 18 metal oxide ENMs, as retrieved from the study of Gajewicz et al. (2015)

Quantum - mechanical descriptors	**Image descriptors**
• Standard enthalpy of formation of metal oxide nanocluster (ΔH_f) • Total energy (TE) • Electronic energy (EE) • Core–core repulsion energy (Core) • Solvent accessible surface (SAS) • Energy of the highest occupier molecular orbital (HOMO) • Energy of the lowest unoccupied molecular orbital (LUMO) • Chemical hardness (η) • Total softness (S) • HOMO-LUMO energy gap (E_g) • Electronic chemical potential (μ) • Valance band (E_v) • Conduction band (E_c) • Mulliken's electronegativity (χ^c) • Parr and Pople's absolute hardness (Hard) • Schuurmann MO shift alpha (Shift) • Polarizability derived from the heat of formation (Ahof) • Polarizability derived from the dipole moment (Ad)	• Volume (V) • Surface diameter (d_S) • Equivalent volume diameter (d_V) • Equivalent volume/surface (d_{Sauter}) • Area (A) • Porosity (P_x) • Porosity (P_y) • Sphericity (Ψ) • Circularity (f_{circ}) • Anisotropy ratio (AR_X) • Anisotropy ratio (AR_Y)

Another dataset that was provided by Zhang et al. (2012) contains information on the toxicity of 24 oxide ENMs: Al_2O_3, CuO, CeO_2, Co_3O_4, CoO, Cr_2O_3, Fe_2O_3, Fe_3O_4, gadolinium oxide (Gd_2O_3), hafnium oxide (HfO_2), In_2O_3, La_2O_3, Mn_2O_3, NiO, Ni_2O_3, Sb_2O_3, SiO_2, SnO_2, R-TiO_2, WO_3, Y_2O_3, ytterbium oxide (Yb_2O_3), ZnO, and ZrO_2 ENMs (data available in the original paper). The toxicity was expressed in terms of $logEC_{50}$, in which EC_{50} means the effective concentration that causes 50% response. The lactate dehydrogenase (LDH), 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium (MTS), and ATP assays were implemented to assess the nanotoxicity to BEAS-2B and rat alveolarmacrophage cells (RAW264.7) cells in the study. Information on the crystalline structure of the ENMs (crystal system, space group, and unit cell parameters), primary and hydrodynamic sizes of metal oxide ENMs, and parameters for calculating ENM band energies (conduction and valence band, band gap energy, absolute electronegativities, and point of zero zeta-potential) were also provided by these authors. Liu et al. (2013b) built a nano-SAR model based on these data along with a summary of the calculated physicochemical properties of the ENMs. Information on 13 descriptors was provided including the ENM primary size (d), energy of conduction band (E_C), energy of valence band (E_V), metal oxide atomization energy (E_{Amz}), metal oxide electronegativity (χ_{MeO}), metal oxide sublimation enthalpy (ΔH_{sub}), metal oxide ionization energy (ΔH_{IE}), metal oxide standard molar enthalpy of formation (ΔH_{sf}), metal oxide lattice enthalpy (ΔH_{Lat}), first molar ionization energy of metal ($\Delta H_{IE,1+}$), ionic index of metal cation (Z^2/r),

IEP, and zeta potential in water at PH of 7.4 (ZP). Data of these descriptors can be accessed in the relevant articles.

3.2.3 Toxicity to *E. coli*

Puzyn et al. (2011) tested the toxicity of 10 metal oxide ENMs to an *E. coli* (Migula) Castellani & Chalmers (ATCC#25254) strain. Metal oxide ENMs covered in the test are Bi_2O_3, CoO, Cr_2O_3, In_2O_3, NiO, Sb_2O_3, SiO_2, V_2O_3, Y_2O_3, and ZrO_2 ENMs. Meanwhile, results of another 7 metal oxide ENMs tested with the same protocol, namely Al_2O_3, CuO, Fe_2O_3, La_2O_3, SnO_2, TiO_2, and ZnO ENMs, were taken from a previous study (Hu et al., 2009) and a dataset consisting of 17 metal oxide ENMs was built. Toxicity to *E. coli* was expressed in terms of the logarithmic values of molar $1/EC_{50}$. Data are shown in Table 3.5.

Meanwhile, information on the characterization of these ENMs in the reported nano-QSARs was presented in light of integrating existing resources and offering reference. As shown in Table 3.5, Kar et al. (2014b) calculated 7 molecular descriptors in their study: metal electronegativity (χ), sum of metal electronegativity for individual metal oxide ($\sum\chi$), sum of metal electronegativity for individual metal oxide divided by the number of oxygen atoms present in a particular metal oxide ($\sum\chi/nO$), N_{Metal}, N_{Oxygen}, the charge of the metal cation corresponding to a given oxide (χ_{ox}), and molecular weight (MW). Two studies (Singh and Gupta, 2014; Toropov et al., 2012) provided 2-dimensional structural information of the ENMs in the form of SMILES (Simplified Molecular Input Line Entry System). Information on ENM size and aggregation size can also be found in Sizochenko's study (Sizochenko et al., 2014). In addition, 12 electronic descriptors were provided (structural parameters of the ENMs were given by Puzyn et al. (2011)), including the standard heat of formation of the oxide cluster (HoF), total energy of the oxide cluster (TE), electronic energy of the oxide cluster (EE), core-core repulsion energy of the oxide cluster (Core), area of the oxide cluster calculated based on COSMO (CA), volume of the oxide cluster calculated based on COSMO (CV), energy of the highest occupier molecular orbital (HOMO) of the oxide cluster, energy of the lowest unoccupied molecular orbital (LUMO) of the oxide cluster, energy difference between HOMO and LUMO energies (GAP), enthalpy of detachment of metal cations Me^{n+} from the cluster surface (ΔH_{Clust}), enthalpy of formation of a gaseous cation (ΔH_{Me+}), and lattice energy of the oxide (ΔH_L). Mu et al. (2016) also presented data of 26 computational descriptors for this dataset, detailed information can be found in the supplemental information of the original publication.

Using the same types of 17 ENMs as in Puzyn's study (Puzyn et al., 2011), Pathakoti et al. (2014) examined the nanotoxicity to the *E. coli* (Migula) Castellani & Chalmers (ATCC#25254) strain under dark conditions and sunlight exposure for 30 minutes. Toxicity

of ENMs was expressed by the logarithmic values of LC50. Information was provided regarding the ENM size (by suppliers), TEM (transmission electron microscopy) particle size, hydrodynamic size, zeta potential in water and in KCl solution, and surface area. Moreover, 6 electronic descriptors for metal oxides and 3 for metal atoms were calculated: the larger (less negative) of the HOMO energies of the alpha spin and beta spin orbitals (HHOMO), the alpha and beta LUMO energies (LUMOA and LUMOB, respectively), the absolute electronegativity of the metal oxide calculated from HHOMO and LUMOA (LZELEHHO), the average of LUMOA and LUMOB (ALZLUMO), molar heat capacity of the metal oxide at 298.15 K (*Cp*), the alpha HOMO and LUMO energies of metal atoms (MHOMOA and MLUMOA, respectively), and the absolute electronegativity of the metal atom calculated from MHOMOA and MLUMOA (QMELECT).

3.3 Existing nano-(Q)SARs

Suitable modeling tools are capable of extracting meaningful relationships between the nano-structures and nanotoxicity, thus yielding predictive models. The widely employed methods concluded from the state-of-the-art of nano-(Q)SARs are linear and logistic regressions, together with the approaches of support vector machines (SVM), artificial neural networks (ANN), and *k*-nearest neighbors (*k*NN) etc. Details on the workflows for model development and the resulting equations (if applicable) are subsequently summarized, including the number of ENMs, predictive performances, descriptors calculation and selection. The datasets used for these nano-(Q)SARs are previously described in Table 3.2. Descriptors used in the developed models or identified factors by relevant studies are summarized in Table 3.6 for further discussion.

3.3.1 Linear regression models

Cellular uptake

In Epa' study (Epa et al., 2012), linear models have been reparameterized for the cell uptake of 108 ENMs (87 in training set, 21 in test set) in PaCa2 and HUVEC cells (Weissleder et al., 2005). A method called multiple linear regression with expectation maximization (MLREM) sparse feature reduction was employed to optimize the descriptor set from a pool of 691 descriptors. DRAGON (v5.5), ADRIANA (v2.2), and an in-house modeling software package were used for descriptor calculation. The best performing models used 19 descriptors for PaCa2 cells ($R^2_{training}$ = 0.76, R^2_{test} = 0.79, *SEE* = 0.19, *SEP* = 0.24) and 11 for HUVEC cells ($R^2_{training}$ = 0.74, R^2_{test} = 0.63, *SEE* = 0.34, *SEP* = 0.36).

Table 3.5. Toxic data to *Escherichia coli* (*E. coli*) reported by Puzyn et al. (2011) and Pathakoti et al. (2014) along with corresponding ENM characterization

Reference	Endpoint/ descriptor[a]	ZrO_2	ZnO	Y_2O_3	V_2O_3	TiO_2	SnO_2	SiO_2	Sb_2O_3	NiO	La_2O_3	In_2O_3	Fe_2O_3	CuO	Cr_2O_3	CoO	Bi_2O_3	Al_2O_3
Puzyn et al., 2011	log $1/EC_{50}$ (mol/L)	2.15	3.45	2.87	3.14	1.74	2.01	2.2	2.64	3.45	2.87	2.81	2.29	3.2	2.51	3.51	2.82	2.49
	HoF (kcal/mol)	-9835	-5307	-11486	-3193	-9826	-2611	-4118	-2141	64	N/A[b]	-3088	-1051	-955	-2829	-8800	-1966	-8244
	TE (au[c])	-23405	-23158	-30634	-26083	-31518	-41962	-21060	-18039	-28053	N/A	-40745	-6971	-45632	-20104	-17007	-36108	-31466
	EE (au)	-358169	-379005	-511019	-441766	-576824	-874369	-321879	-221602	-432596	N/A	-872315	-44000	-874569	-307815	-298812	-695663	-630309
	Core (au)	334764	355847	480385	415683	545306	832407	300818	203563	404543	N/A	831570	37029	828937	287711	281806	659555	598843
	CA (A^2)	1055	855	1805	1130	1100	1734	753	975	659	N/A	1314	243	639	659	1072	1551	1109
	CV (A^3)	2403	1849	5401	2426	2340	3959	1467	1797	1088	N/A	3095	319	1108	1161	1548	4107	2260
	HOMO (eV)	-6.19	-10.76	-1.28	-3.54	-10.33	-6.14	-7.12	-8.3	-5.75	N/A	-8.16	-7.13	-6.11	-6.9	-10.48	-4.11	-4.88
	LUMO (eV)	-4.54	-6.89	1.2	0.64	-2.86	-2.29	-3.89	-1.03	-1.03	N/A	-3.37	-0.68	-2.25	-0.49	-8.28	-1.4	-0.29
	GAP (eV)	-1.65	3.87	-2.48	-4.17	-7.47	-3.85	-3.23	-7.27	-4.73	N/A	-4.79	-6.45	-3.85	-6.41	-2.2	-2.71	-4.59
	ΔH_{Clust} (kcal/mol)	-8956	-5357	-11485	-3157	-8731	-2091	-3295	-1526	325	N/A	-3190	-140	-759	-2264	-8318	-1601	-8017
	ΔH_{Me+} (kcal/mol)	1358	662	837	1098	1576	1717	1686	1233	597	1017	1271	1408	706	1269	602	1137	1188
	ΔH_L (kcal/mol)	-2641	-971	-3111	-3555	-2896	-2821	-3158	-3281	-965	-2969	-3449	-3589	-992	-3645	-933	-3199	-3695

Table 3.5.(Continued)

Pathakoti et al., 2014	toxicity under darkness, log $1/EC_{50}$ (mol/L)	2.58	5.8	5.79	3.48	2.14	2.53	2.54	3.12	3.79	4.96	2.83	2.4	4.24	2.06	3.13	3.55	2.42
	toxicity under sunlight exposure, log $1/EC_{50}$ (mol/L)	3.04	6.23	5.84	3.78	4.68	3.24	2.92	3.66	3.87	5.56	3.48	2.54	5.71	2.06	3.33	4.02	2.75
	Particle size (vendor) (nm)	<100	<100	<50	N/A	<100	<100	10–20	90–210	<50	<100	<100	<50	<50	<100	<100	90–210	<50
	Particle size TEM (nm)	27±6	71±17	38±9	N/A	42±9	15±4	20±5	84±23	14±9	65±19	60±14	68±20	28±7	47±27	55±13	144±7	55±17
	Hydrodynamic size (nm)	2337	1614	357	307	748	3971	1230	619	399	508	308	>6000	285	426	262	4084	330
	Zeta potential (mV) (H_2O)	−(6.9 ±0.5)	−(20.9 ±0.5)	16.3± 0.9	−(27.9 ±0.9)	−(10.7 ±2.5)	−21.1 ±0.4	−29.8 ±1.9	−20.7 ±1.3	26.0± 0.4	−(3.6 ±1.1)	22.6± 0.4	−(6.3 ±1.0)	24.4± 0.6	−(12.0 ±1.3)	17.5± 1.5	−(16.5 ±0.8)	30.3± 1.3
	Zeta potential (mV) (KCl)	4.0±2.7	−(24.9 ±0.3)	17.9± 1.0	−(32.6 ±0.5)	−(2.2 ±0.4)	−(16.7 ±0.2)	−(33.7 ±1.6)	−(12.7 ±0.4)	26.8± 1.2	22.3± 1.7	28.7± 0.4	−(19.5 ±1.9)	19.1± 0.3	23.3± 1.0	26.0± 0.5	−(4.9 ±0.1)	25.3± 1.1
	Surface area (m^2/g)	22	15	31	N/A	36	18.6	N/A	N/A	80	20	28	36	33	N/A	>8	N/A	37
	HHOMO (au)	-0.2430	-0.2282	-0.1894	-0.2192	-0.2645	-0.3053	-0.3432	-0.2624	-0.2412	-0.1872	-0.2646	-0.2833	-0.2358	-0.2451	-0.2213	-0.2528	-0.2833
	LZELEHHO (au)	0.1842	0.1320	0.1285	0.1739	0.1949	0.2242	0.2450	0.1740	0.1804	0.1205	0.1960	0.1747	0.1782	0.1985	0.1685	0.1845	0.2105
	LUMOA (au)	-0.1255	-0.0358	-0.0676	-0.1286	-0.1252	-0.1431	-0.1467	-0.0855	-0.1197	-0.0539	-0.1274	-0.0661	-0.1206	-0.1519	-0.1158	-0.1161	-0.1377
	LUMOB (au)	-0.1255	-0.1387	-0.0676	-0.1058	-0.1252	-0.1431	-0.1467	-0.0855	-0.1138	-0.0539	-0.1274	-0.1625	-0.1189	-0.1172	-0.1307	-0.1161	-0.1377
	ALZLUMO (au)	-0.1255	-0.0872	-0.0676	-0.1172	-0.1252	-0.1431	-0.1467	-0.0855	-0.1167	-0.0539	-0.1274	-0.1143	-0.1197	-0.1345	-0.1233	-0.1161	-0.1377
	Cp ($J\ mol^{-1}\ K^{-1}$)	56.19	40.25	102.51	103.22	55.48	52.59	44.43	101.63	44.31	108.78	92	103.85	42.3	118.74	55.23	113.51	79.04
	MHOMOA (au)	-0.2318	-0.2925	-0.2109	-0.2468	-0.2324	-0.2670	-0.3014	-0.3343	-0.2363	-0.1881	-0.2016	-0.2294	-0.2887	-0.2222	-0.2324	-0.3190	-0.2185
	MLUMOA (au)	0.0164	0.0434	0.0177	0.0240	0.0214	-0.0168	-0.0065	0.1302	0.0352	0.0152	0.0105	0.0311	0.0365	0.0273	0.0355	0.1137	0.0170
	QMELECT (au)	0.1077	0.1246	0.0966	0.1114	0.1055	0.1419	0.1539	0.1021	0.1006	0.0864	0.0956	0.0992	0.1261	0.0975	0.0984	0.1027	0.1008

Table 3.5.(Continued)

Kar et al., 2014b	χ	1.33	1.65	1.22	1.63	1.54	1.96	1.9	2.05	1.91	1.1	1.78	1.83	1.9	1.66	1.88	2.02	1.61
	$\sum\chi$	1.33	1.65	2.44	3.26	1.54	1.96	1.9	4.1	1.91	2.2	3.56	3.66	1.9	3.32	1.88	4.04	3.22
	$\sum\chi$/nO	0.665	1.65	0.813	1.087	0.77	0.98	0.95	1.367	1.91	0.733	1.187	1.22	1.9	1.107	1.88	1.347	1.073
	MW	123.2	81.38	225.82	149.88	79.86	150.7	60.08	291.52	74.69	325.8	277.62	159.6	79.546	151.98	74.93	465.96	101.96
	N_{Metal}	1	1	2	2	1	1	1	2	1	2	2	2	1	2	1	2	2
	N_{Oxygen}	2	1	3	3	2	2	2	3	1	3	3	3	1	3	1	3	3
	χ_{av}	4	2	3	3	4	4	4	3	2	3	3	3	2	3	2	3	3
Singh and Gupta, 2014; Toropov et al., 2012	SMILES notation	O = [Zr] = O	O = [Zn]	O = [Y]O[Y] = O	O = [V]O[V] = O	O = [Ti] = O	O = [Sn] = O	O = [Si] = O	O = [Sb]O[Sb] = O	[Ni] = O	O = [La]O[La] = O	O = [In]O[In] = O	O = [Fe]O[Fe] = O	[Cu] = O	O = [Cr]O[Cr] = O	[Co] = O	O = [Bi]O[Bi] = O	O = [Al]O[Al] = O
Sizochenko et al., 2014	Size (nm)	47	71	38	15	46	15	150	20	30	46	30	32	N/A	60	100	90	44
	Aggregation size (nm)	661	189	1223	1307	265	810	640	223	291	673	224	298	N/A	617	257	2029	372

[a]EC50 - the effective concentration that causes 50% response; HoF[2] - the standard heat of formation of the oxide cluster; TE - total energy of the oxide cluster; EE - electronic energy of the oxide cluster; Core - core-core repulsion energy of the oxide cluster; CA - area of the oxide cluster calculated based on COSMO; CV - volume of the oxide cluster calculated based on COSMO; HOMO - energy of the highest occupier molecular orbital of the oxide cluster; LUMO - energy of the lowest unoccupied molecular orbital of the oxide cluster; GAP - energy difference between HOMO and LUMO energies; ΔH_{Clust} - enthalpy of detachment of metal cations Me^{n+} from the cluster surface; ΔH_{Me+}- enthalpy of formation of a gaseous cation; ΔH_L - lattice energy of the oxide; [b]N/A – data not available; [c]au - atomic units

A partial least squares (PLS) model predicting the cellular uptake ($\log_{10}$[ENM]/cell pM) of 109 magnetofluorescent ENMs in PaCa2 cells (Weissleder et al., 2005) was constructed by Kar et al. (2014a). In this study, a set of 307 descriptors was calculated using the Cerius 2 (v4.10), DRAGON (v6), and PaDEL-Descriptor (v2.11) which was afterwards filtered by the genetic function approximation (GFA). Finally, six molecular descriptors appeared in the developed model:

$$log_{10}[NP]/cell = 3.335 + (0.774 \times < 1 - Atype - N - 66 >) - (0.222 \times Atype - N - 67) + \left(7.360 \times < 0.600 - \sum \beta' >\right) - (0.101 \times Jurs - RPCS) - (0.00002 \times Wap) - (0.462 \times nRNO2)$$

$n_{\text{training}} = 89$, $LV = 5$, $R^2 = 0.806$, $Q^2_{LOO} = 0.758$, $Q^2_{Leave\text{-}10\%\text{-}out} = 0.634$, $Q^2_{Leave\text{-}25\%\text{-}out} = 0.648$, $SEE = 0.20$, $\overline{r^2_{m(LOO)Scaled}} = 0.665$, $\Delta r^2_{m(LOO)Scaled} = 0.113$, $n_{\text{test}} = 20$, $Q^2_{F1} = R^2_{pred} = 0.879$, $SEP = 0.12$,

$Q^2_{F2} = 0.868$, $\overline{r^2_{m(test)Scaled}} = 0.793$, $\Delta r^2_{m(test)Scaled} = 0.115$,

$\overline{r^2_{m(overall)Scaled}} = 0.679$, $\Delta r^2_{m(overall)Scaled} = 0.116$

In the model, the descriptors *Atype* - N - 66 and *Atype* - N - 67 are the hydrophobicity of the N atom in respectively a primary and a secondary aliphatic amine (Al-NH_2 and Al_2-NH, respectively), $\sum \beta'$ characterizes the measure of electronic features of the molecule relative to molecular size, *Jurs–RPCS* stands for the relative positive charge surface area, *Wap* represents for the all-path Wiener index, and *nRNO*2 is the number of aliphatic nitro groups. The leverage and distance to model in X-space (DModX) approaches (Gramatica, 2007; Wold et al., 2001) was applied to check model's domain of applicability.

Using the same data from Weissleder et al. (2005), Ghorbanzadeh et al. (2012) proposed a predictive model of cellular uptake ($\log_{10}$[ENM]/cell pM) on the basis of a multilayered perceptron neural network technique. A self-organizing map (SOM) strategy was employed combined with stepwise MLR to promote the feature reduction. This procedure provided six most informative descriptors, namely number of donor atoms (N and O) for H-bonds (nHDon), Geary autocorrelation of lag 1 weighted by van der Waals volume (GATS1v), 3D-MoRSE-signal 29/unweighted (Mor29u), D total accessibility index/weighted by Sanderson electronegativity (De), 3D-MoRSE-signal 14/unweighted (Mor14u), as well as the mean electrotopological state (Ms). The linear model has the form:

$$log_{10}[NP]/cell = 2.970 - 0.130 \times \text{nHDon} + 0.412 \times \text{GATS1v} - 0.398 \times \text{Mor29u} + 1.243 \times \text{De} - 0.163 \times \text{Mor14u} + 0.045 \times \text{Ms}$$

The model gave a correlation coefficient (R) of 0.782 for the training set (RMSE = 0.369) and 0.755 for the prediction (RMSE = 0.357). Williams plot was subsequently put into use for visualizing the domain of model's applicability.

Cytotoxicity

Based on the apoptosis assay of smooth muscle cells from Shaw et al. (2008), Epa et al. (2012) developed a model consisting of three descriptors for the core material (I_{Fe3O4}), surface coating ($I_{dextran}$), and surface charge ($I_{surf.chg}$) of ENMs. The descriptors are considered to have a value of 1 when the condition is present, and 0 when the condition is absent. For instance, I_{Fe3O4} is set to be 1 for the ENM with Fe_2O_3 core, and 0 when the ENM core is Fe_3O_4; $I_{dextran}$ is equal to 1 in case of a dextran coating and 0 for others; surface functionality is encoded as 1 (basic), −1 (acidic), or 0 (neutral). Smooth muscle apoptosis was used as the endpoint in the constructed model:

$$SMA = 2.26(\pm 0.72) - 10.73(\pm 1.05) \times I_{Fe_2O_3} - 5.57(\pm 0.98) \times I_{dextran} - 3.53(\pm 0.54) \times I_{surf.chg}$$

where n = 31, $R^2_{training}$ = 0.81, R^2_{test} = 0.86, SEE = 3.6, SEP = 3.3.

Papa et al. (2015) reported three MLR models predicting the potential of ZnO and TiO_2 ENMs inducing the release of LDH in rat lung cells. Data was retrieved from the study of Sayes and Ivanov (2010) which provided values of five descriptors including engineered size (X0), size in water (X1), size in phosphate buffered saline (X2), concentration (X4), and zeta potential (X5). The first linear model combined information on both TiO_2 and ZnO ENMs (all together 31 ENMs):

$$LDH_{(TiO_2+ZnO)} = 0.66 + 0.003X4 + 0.005X0 - 4.46E - 5X2$$

R^2 = 0.82, Q^2_{loo} = 0.76, $Q^2_{lmo30\%}$ = 0.74, r^2_{YS} = 0.10, s = 0.11, F= 40. The Williams plot for applicability domain of the model was depicted in the original publication. Besides, linear models were also built separately for TiO_2 (22 ENMs) and ZnO ENMs (15 ENMs):

$$LDH_{(TiO_2)} = 0.599 + 0.003X4 + 0.004X0$$

R^2 = 0.84, Q^2_{loo} = 0.79, $Q^2_{lmo30\%}$ = 0.78, r^2_{YS} = 0.10, s = 0.12, F = 48

$$LDH_{(ZnO)} = 1.041 + 0.001X1 - 0.001X2 + 0.001X4$$

R^2 = 0.91, Q^2_{loo} = 0.80, $Q^2_{lmo30\%}$ = 0.76, r^2_{YS} = 0.22, s = 0.08, F = 35.

Another approach explicitly and completely based on MLR is reported by Gajewicz et al. (2015). In this case, the cytotoxicity of 18 metal oxide ENMs to the HaCaT cell line was modeled. A set of 27 descriptors were calculated including 16 quantum-mechanical descriptors and 11 image descriptors derived from Transmission Electron Microscopy (TEM) images. For calculating the quantum-mechanical descriptors, the molecular geometry was optimized at the level of the semi-empirical PM6 method (Stewart, 2007) encoded in MOPAC 2009 (Stewart, 2009). Information on the size, size distribution, shape, porosity, and surface area of ENMs was extracted based on TEM images to generate the 11 image descriptors. Two descriptors were afterwards selected by the genetic algorithm (GA), i.e., ΔH_f^c and χ^c. The model can be expressed as:

$$\log(LC_{50})^{-1} = 2.47(\pm 0.05) + 0.24(\pm 0.05) \times \Delta H_f{}^c + 0.39(\pm 0.05) \times \chi^c$$

F = 44.6, $p = lx10^{-4}$, n = 18, R^2 = 0.93, $RMSE_C$ = 0.12, Q^2_{CV} = 0.86, $RMSE_{CV}$ = 0.16, Q^2_{EXT} = 0.83, $RMSE_p$ = 0.13

where ΔH_f^c is the enthalpy of formation of metal oxide nanocluster representing a fragment of the surface and χ^c represents the Mulliken's electronegativity of the cluster. The domain of applicability of the model was described by means of a Williams plot.

Using the dataset reported by Gajewicz et al. (2015), Pan et al. (2016) developed two predictive models incorporating the so-called Improved SMILES-Based Optimal Descriptors. The models predicting the cytotoxicity of metal oxide ENMs to HaCaT cells have the forms:

$$\log\left(\frac{1}{\mathrm{LC}_{50}}\right) = -0.2909(\pm 0.0664) + 0.1038(\pm 0.0027) \times \mathrm{DCW}(1,3)$$

n = 13, R^2 = 0.9606, Q^2_{LMO} = 0.9393, s = 0.008, F = 268, p < 0.0001; and

$$\log\left(\frac{1}{\mathrm{LC}_{50}}\right) = 0.0012(\pm 0.0048) + 0.0778(\pm 0.0001) \times \mathrm{DCW}(1,3)$$

n = 12, R^2 = 0.9997, Q^2_{LMO} = 0.9996, s = 0.007, F = 1273, p < 0.0001. The number 1 in DCW(1,3) is the coefficient for classification of features into two classes (noise and active); the number 3 in DCW(1,3) is the number of epochs of the Monte Carlo optimization. The

characteristics of ENMs involved in the models are namely molecular weight, cationic charge, mass percentage of metal elements, individual size, and aggregation size of ENMs.

Besides, in the study of Liu et al. (2013b) a linear regression model was developed for 24 metal oxide ENMs based on a recently reported dataset (Zhang et al., 2012). Three descriptors were involved in the model, namely E_c, ΔH_{IE}, and χ_{MeO}. The model was reported to give an accuracy of 89% for the samples.

Toxicity to E. coli

Puzyn et al. (2011) originally built a dataset for the toxicity of 17 metal oxide ENMs to *E. coli*. Based on the data, a simple and statistically significant nano-QSAR model was obtained which used a single descriptor ΔH_{Me+}:

$$\log(1/EC_{50}) = 2.59 - 0.50 \times \Delta H_{Me+}$$

$R^2 = 0.85$, $RMSE_C = 0.20$, $Q^2_{CV} = 0.77$, $RMSE_{CV} = 0.24$, $Q^2_{EXT} = 0.83$, $RMSE_P = 0.19$

Calculation of a pool of 12 variables (Table 3.5) was executed using the PM6 method as implemented in MOPAC 2009. GA was applied for selecting the most informative descriptors. PLS Toolbox and the Statistics Toolbox for MATLAB were utilized for model development. The leverage approach and Williams plot were employed to visualize model applicability domain.

Working on the same dataset from Puzyn et al. (2011), Kar et al. (2014b) built a stepwise MLR model as well as a PLS model. Seven descriptor were used for model construction namely χ, $\Sigma\chi$, $\Sigma\chi/nO$, N_{Metal}, N_{Oxygen}, χ_{ox}, and MW (Table 3.5). For the MLR model feature reduction was accomplished by the 'stepping criteria' (F), and only the descriptor χ_{ox} was seen in the model:

$$\log(1/EC_{50}) = 4.781 - (1.380 \times \chi_{ox})$$

$n = 17$, $R^2 = 0.84$, $R^2_{adj} = 0.83$, $Q^2_{LOO} = 0.81$, $Q^2_{L\text{-}10percent\text{-}OUT} = 0.82$,
$Q^2_{L\text{-}20percent\text{-}OUT} = 0.83$, $Q^2_{L\text{-}25percent\text{-}OUT} = 0.80$, $^cR^2_P = 0.82$

Meanwhile the developed PLS model contained two descriptors χ_{ox} and χ, and has the form:

$$\log(1/EC_{50}) = 4.401 - (1.324 \times \chi_{ox}) + (0.176 \times \chi)$$

$n = 17$, $LV = 1$, $R^2 = 0.82$, $Q^2_{LOO} = 0.75$, $Q^2_{L\text{-}10\text{percent-OUT}} = 0.76$,
$Q^2_{L\text{-}20\text{percent-OUT}} = 0.74$, $Q^2_{L\text{-}25\text{percent-OUT}} = 0.76$, ${}^{c}R^2_P = 0.79$

Characterization of the applicability domain of the model was performed by the leverage approach (Gramatica, 2007).

Mu et al. (2016) also reported MLR models building on the data from Puzyn et al. (2011). Calculation of descriptor was performed using PM6 methods within MOPAC 2012 software package. Approaches of Pearson and pair-wise correlations, and clustering and principal component analysis were incorporated to obtain optimal structure descriptors for modeling. Among the developed models, a simple but statistically significant nano-QSAR has the form:

$$\log\left(\frac{1}{\mathrm{EC}_{50}}\right) = (4.412 \pm 0.165) + (-0.121 \pm 0.068)Z/r + (-0.001 \pm 2.57 \times 10^{-4})\Delta H_{\mathrm{Me+}}$$

where Z is the ionic charge, r is the Pauling ionic radius. Statistical indicators of the model are: R^2 = 0.8793, $RMSE$ = 0.442, F = 55.654, p = 4.23 × 10^{-7}. Leverage approach and Williams plots were used for the characterization of model applicability domain. Based on the developed model, toxic potencies of other 35 metal oxide ENMs were predicted and visualized in a periodic table. Other models using different descriptors were also described in the study.

Pan et al. (2016) also built *in silico* models using data from Puzyn et al. (2011). The reported models on the basis of the Improved SMILES-Based Optimal Descriptors can be expresses as:

$$\log\left(\frac{1}{\mathrm{LC}_{50}}\right) = 0.0321(\pm 0.1443) + 0.2658(\pm 0.0141) \times \mathrm{DCW}(6,11)$$

n = 10, R^2 = 0.8891, Q^2_{LMO} = 0.8378, s = 0.179, F = 164, p < 0.0001; and

$$\log\left(\frac{1}{\mathrm{LC}_{50}}\right) = -0.0076(\pm 0.0306) + 0.1420(\pm 0.0020) \times \mathrm{DCW}(6,17)$$

n = 9, R^2 = 0.9824, Q^2_{LMO} = 0.9745, s = 0.007, F = 391, p < 0.0001. The characteristics of ENMs involved in the models are namely molecular weight, cationic charge, mass percentage of metal elements, individual size, and aggregation size of ENMs.

3.3.2 Logistic regression models

Liu et al. (2011) constructed logistic regression models to classify the effect of nine metal oxide ENMs to BEAS-2B cells into toxic (T) or nontoxic (N). The model with the best classification performance is:

$$\ln\left(\frac{P(N\ P \in T)}{P(N\ P \in N)}\right) = 3600.6 + 103.5 \times d + 9.5 \times \theta_v + 97.6 \times P_{\mathrm{Me}} - 58.5 \times E_{\mathrm{MeO}}$$

where $P(N\ P \in T)$ and $P(N\ P \in N)$ are the probabilities of an ENM being classified as toxic or nontoxic, respectively. d is the size of ENM; θ_v is the volume concentration derived from the mass concentration of ENMs; P_{Me} is the period of the ENM metal in the periodic table; E_{MeO} is the atomization energy of the metal oxide. Model applicability domain was depicted by the principal component analysis.

Liu et al. (2013b) developed two nano-SAR models based on the logistic regression and quadratic logistic regression methods, respectively. The dataset of Zhang et al. (2012) was chosen. This dataset covered data on the toxicity of 24 metal oxide ENMs to BEAS-2B and RAW264.7 cell lines as described above. The quadratic logistic regression model was shown to achieve an accuracy of 89.97 % with only two descriptors E_C and Z^2/r. Meanwhile a marginally better predictability of 90.09% for the logistic regression model was obtained. The molecular descriptors that were included in the logistic regression model were E_C, $E_{\Lambda mz}$, and d.

Logistic regression models were also built by Liu et al. (2013a) based on an integration of multiparametric bioactivity assays of 44 iron oxide ENMs (Shaw et al., 2008). The conception of 'hit' (significant bioactivity, Signal-to-Noise Ratio > 1.645) was utilized in the study, and the number of hits served as the bioactivity class definition (identifying an ENM as bioactive or inactive) enabling nano-SAR development. Clustering analysis via SOM was also considered besides the number of hits as an alternative to define a class. ENM descriptors included the primary size, zeta potential, R1 and R2. Results showed that the logistic regression model based on class definition of H5 (five hits) possesses the best predictability of 79.3 %, using ENM size and R2 as descriptors. The class definition H6 also enabled the construction of a simple logistic regression model (R1 as the sole descriptor) with 78.2% accuracy.

3.3.3 Support vector machine models

A SVM classification model has been developed by Fourches et al. (2010) using the experimental data of 44 ENMs from Shaw et al. (2008). ENM size, R1, R2, and zeta potential were used as input descriptors, and an arbitrary threshold at Z_{mean} = -0.40 was applied to enable a binary classification. Three clusters of ENMs were identified after assigning a hierarchical clustering procedure. It was found that all monocrystalline iron oxide ENMs were in cluster II and all the quantum dots appeared in the cluster I. Results of classification confirmed the good predictability of the clustering-based nano-SARs (5-fold external cross-validation) in the cluster II:

Cluster I: n = 13, *sensitivity* = 0.5, *specificity* = 0.8;

Cluster II: n = 18, *sensitivity* = 0.78, *specificity* = 0.78;

Cluster III: n = 13, *sensitivity* = 0.7, *specificity* = 0.4

where sensitivity = (number of true positives)/(total number of true positives), and specificity = (number of true negatives)/(total number of true negatives) for the binary classification problems.

Another SVM nano-SAR classifying 23 metal oxide ENMs as toxic or nontoxic was built by Liu et al. (2013b), based on measured toxicological responses in BEAS-2B cells and murine myeloid RAW 264.7 cells following an established protocol (Zhang et al., 2012). A SOM based consensus clustering was employed and afterwards identified three ENM clusters. The clusters II and III contained ENMs being reported as toxic, and thus were grouped into a single cluster of ENMs classified as having a positive response. ENMs in cluster I were labeled as nontoxic. A pool of 30 descriptors were initially considered including information on the fundamental metal oxide, energies or enthalpies of metal oxide, ENMs size, zeta potential and isoelectric point, and ENM energy. Descriptor selection was accomplished by the evaluation of models derived from all possible descriptor combinations. The SVM algorithm successfully correlated the cytotoxicity of ENMs with ENM conduction band energy (E_C) and ionic index of metal cation (Z^2/r). The penalty factor and the kernel width of the SVM model were determined to be 128 and 2, respectively. The discriminant function of the SVM model was given by

$$f(\mathbf{x}) = \sum_{i=1}^{6} \alpha_i e^{-2[(x_{i,1}-x_1)^2+(x_{i,2}-x_2)^2]} + b$$

where **x** refer to the ENM identified by the normalized descriptors vector [Z^2/r, E_C] (i.e., x_1, x_2), $x_{i,1}$ and $x_{i,2}$ stand for the normalized first and second descriptors identified as support

vectors. The values of α_i (i = 1-6) were represented by ZnO (82.342), Ni_2O_3 (128), Mn_2O_3 (83.696), NiO (-70.471), CeO_2 (-95.566), and Fe_2O_3 (128) with b being -10.888. The model was reported to give a predictive accuracy (obtained via 0.632 estimator) of 93.74%. Model applicability domain was characterized by a probabilistic approach (Netzeva et al., 2005).

3.3.4 Artificial neural network models

Based on the experimental results of Shaw et al. (2008), a Bayesian regularized ANN model was constructed predicting the smooth muscle cells' apoptosis triggered by 31 ENMs (Epa et al., 2012). Model statistics were as follows: n = 31, $R^2_{training}$ = 0.80, R^2_{test} = 0.90, SEE = 2.8, SEP = 2.9. Meanwhile an ANN nano-SAR was also built in the study, modeling the cellular uptake in HUVEC and PaCa2 cells (Weissleder et al., 2005):

cellular uptake in HUVEC cells: $R^2_{training}$ = 0.70, SEE = 0.30, R^2_{test} = 0.66, SEP = 0.33, descriptor number = 11;

cellular uptake in PaCa2 cells: $R^2_{training}$ = 0.77, SEE = 0.15, R^2_{test} = 0.54, SEP = 0.28, descriptor number = 19.

Besides the above-mentioned MLR model developed by Ghorbanzadeh et al. (2012), another nano-SAR on the basis of a multilayered perceptron neural network technique was also introduced in their study. The SOM strategy combined with stepwise MLR selected six most informative descriptors namely nHDon, GATS1v, Mor29u, De, Mor14u, and Ms. The derived model gave a performance in terms of values of R^2 of 0.934 for the training set, 0.945 for the internal test set, and 0.943 for the external test set. The calculated RMSE values are 0.146, 0.121, and 0.214 for respective training, internal, and external test sets, while the corresponding values of F are 531, 142, and 65, respectively. The applicability domain of the model was firstly evaluated by the approach based on ranges of individual descriptors. A Williams plot was subsequently put into use for visualizing the domain of applicability.

3.3.5 k-nearest neighbor models

A classification model employing the kNN approach was developed in the study of Fourches et al. (2010). The model was proposed to predict the cellular uptakes of 109 ENMs in PaCa2 cells (Weissleder et al., 2005). Coefficients of correlation R_{abs}^2 were shown to range from 0.65 to 0.80 for the external sets, and from 0.67 to 0.90 taking into account the applicability domain which was defined by the Euclidean distance approach. In the study, the descriptors were identified that most frequently occurred in the models (1-5 fold

cross-validations) with the highest prediction accuracy. The top 10 descriptors ranked by averaged frequency were reported to be SlogP_VSA1, SlogP_VSA2, SlogP_VSA5, b_double, SlogP_VSA0, PEOE_VSA+1, vsa_don, vsa_other, vsa_base, and PEOE_VSA_FPOS. The SlogP_VSA0 and SlogP_VSA1, along with other descriptors with relatively low frequency such as GCUT_SLOGP_0 and BCUT_SLOGP_0, are considered to be generally related to the lipophilicity. For instance, the PaCa2 uptake of ENMs was observed to be positively correlated with the enrichment of lipophilic compounds on ENM surfaces (value of GCUT_SLOP_0). Other discriminated factors affecting the PaCa2 uptake of ENMs were found to be about the molecular refractivity, the specific van der Waals surface area, and the electrostatic properties. The applicability domain of the model was characterized by the Euclidean distance.

An attempt of predicting the cytotoxicity of 44 iron oxide ENMs based on *k*NN was also reported by Liu et al. (2013a). As described above, different numbers of hits were discussed in the study for introducing class definitions besides the clustering analysis via SOM. The results showed that a *k*NN model using SOM-based consensus clustering gave the best predictive performance of 74.9 % accuracy. Three descriptors, ENM size, R1, and R2 were obtained in this model. Meanwhile, H4 class definition was also deemed to be a suitable choice which enabled the development of a *k*NN model correctly predicting 74.3% of the samples.

3.3.6 Other models

Chau and Yap (2012) attempted to correlate the cellular uptake in PaCa2 with the calculated parameters from PaDEL-Descriptor (v2.8). By lowering the threshold value of being significant uptake into PaCa2, 56 ENMs with cellular uptake of more than 5000 ENMs per cell (Weissleder et al., 2005) were considered as a positive class, and the other 49 were defined as the negative class. Based on the four modeling techniques of naive Bayesian classifier (NBC), logistic regression, *k*NN, and SVM, 2100 candidate models were developed while only 102 of them were qualified according to the selection criteria. To build a final consensus nano-SAR model, the top 5 candidate models were chosen consisting of 3 *k*NN, 1 SVM, and 1 NB models. The consensus model gave a good predictive performance with sensitivity of 98.2% and specificity of 76.6% for the dataset. Descriptors that commonly appeared in the candidate models were number of CH_2 groups, primary, secondary and tertiary nitrogens, halogens (fluorine, bromine, iodine), sulphur atoms, fused rings and hydrogen bonding. Most of the descriptors that contributed to the model were interpreted as related to the lipophilicity (e.g., number of lipophilic groups). Other factors such as the hydrogen bonding between nitrogen and hydrogen, and the sulphur and various

halogen atoms were also found to affect the cellular uptake of ENMs. This is in agreement with the study of Fourches et al. (2010).

Chen et al. (2016) reported several nano-SARs for the categorization of ENM hazards to different biota. The toxicity data was retrieved from the database of Chen et al. (2015) and the online chemical modeling environment platform (Sushko et al., 2011). Approaches of functional tree, C4.5 decision tree, random tree, and Simple CART were employed for model development. Global nano-SARs across species using LC50 data were shown to correctly predict more than 70% of the samples in training (320 ENMs) and test sets (80 ENMs) based on functional tree, C4.5 decision tree, random tree methods. The species-specific nano-SARs were also derived for *Danio rerio*, *Daphnia magna*, *Pseudokirchneriella subcapitata*, and *Staphylococcus aureus* with good predictivity. Summarized from the developed models, the molecular polarizability, accessible surface area, and solubility were identified as key factors affecting the biological activities of metallic ENMs.

Moreover, Zhang et al. (2012) reported a regression tree model using the metal dissolution of metal oxide ENMs and energy of conduction band to predict the toxicity potential of 24 metal oxide ENMs. With the data from Zhang et al. (2012), Sizochenko et al. (2015) developed causal inference nano-SARs for BEAS-2B and RAW 264.7 cell lines (24 metal oxide ENMs for each cell line) with high predictivity. Luan et al. (2014) and Kleandrova et al. (2014) developed the novel QSTR-perturbation (quantitative structure–toxicity relationship) models assessing the cytotoxicity and ecotoxicity of various types of ENMs. The factors of molar volume, polarizability, size of ENMs, electronegativity, and the hydrophobicity and polar surface area of surface coatings were indicated by the reported models. Singh and Gupta (2014) previously performed three cases of nano-(Q)SAR study for metallic ENMs on the basis of the datasets generated by Puzyn et al. (2011), Shaw et al. (2008), and Weissleder et al. (2005). In the study, classification and regression models were constructed predicting various biological effects of the ENMs by an ensemble learning based strategy called stochastic gradient boosting and bagging algorithms. Results showed that the developed models are of robustness and no over-fitting of data was present in all case studies. Besides, attempts to link the information of ENM structures to corresponding biological effects were also made using other modeling techniques, such as Monte-Carlo method (Toropov et al., 2012; 2013), NBC and linear discriminate analysis (Liu et al., 2013a,b), random forest regression (Sizochenko et al., 2014; 2015), and self-written least-squares fitting program (Pathakoti et al., 2014).

Table 3.6. Overview of computational descriptors or factors discussed in nano-(Q)SAR studies, including information on the original dataset for modeling. Name of the descriptors in original publications are given in the parenthesis (if available)

Reference	Descriptor or identified factor by developed models	Dataset
Chau and Yap, 2012	Number of CH_2 groups, primary, secondary and tertiary nitrogen, halogens (fluorine, bromine, iodine), sulphur atoms, fused rings, hydrogen bonding	Weissleder et al., 2005
Epa et al., 2012	Number of 10 membered rings (nR10), molecular asphericity (ASP), d COMMA2 value/weighted by atomic masses (DISPm), Qzz COMMA2 value/weighted by atomic masses (QZZm), number of secondary amides, aliphatic (nRCONHR), number of (thio-) carbamates, aromatic (nArOCON), CH3X (C-005), number of circuits (nCIR), number of N atoms (nN), average molecular span R (SPAM), Qyy COMMA2 value/weighted by atomic polarizabilities (QYYp), number of total secondary C sp^3 (nCs), number of aromatic hydroxyls (nArOH), H attached to C0(sp^3) with 2X attached to next C (H-053), =O (O-058)	
Fourches et al., 2010	Surface area "owned" with SlogP weight −10 to −0.40 (SlogP_VSA0), surface area "owned" with SlogP weight −0.40 to −0.20 (SlogP_VSA1), surface area "owned" with SlogP weight −0.20 to 0 (SlogP_VSA2), surface area "owned" with SlogP weight −0.15 to −0.20 (SlogP_VSA5), van der Waals surface area surface area of hydrogen-bond donors (vsa_don), van der Waals surface area of nondonor/-acceptor atoms (vsa_other), van der Waals surface area surface area of basic atoms (vsa_base), sum of the van der Waals surface area of atoms whose PEOE partial charge is positive, divided by the total surface area (PEOE_VSA_FPOS), van der Waals surface area where atomic partial charge 0.05<q<0.10 (PEOE_VSA+1), number of double bonds, aromatic bonds are not considered (b_double)	
Ghorbanzadeh et al., 2012	Number of donor atoms for H-bonds (nHDon), Geary autocorrelation of lag 1 weighted by van der Waals volume (GATS1v), 3D-MoRSE-signal 29/unweighted (Mor29u), D total accessibility index/weighted by Sanderson electronegativity (De), 3D-MoRSE-signal 14/unweighted (Mor14u), mean electrotopological state (Ms)	
Kar et al., 2014a	Hydrophobicity of the N atom in primary aliphatic amine (Al-NH_2) fragment (*Atype* – N – 66), hydrophobicity of the N atom in a secondary aliphatic amine (Al_2-NH) fragment (*Atype* – N – 67), measure of electronic features of the molecule relative to molecular size (∑β′), relative positive charge surface area (*Jurs–RPCS*), all-path Wiener index (*Wap*), number of aliphatic nitro groups (*nRNO2*)	
Singh and Gupta, 2014	Weighted partial negative surface area-3 (WNSA-3), weighted partial positive area-2 (WPSA-2), Chi simple path descriptor of order 5 (SP-5), Chi valance path descriptor of order 4 (VP-4), moment of inertia along X/Z-axis (MOMI-XZ), logarithmic form of octanol–water partition coefficient predicted by atomic method (XlogP), number of rotatable bonds (nRotB), number of hydrogen bond donors (nHBDon), Chi valance path cluster of order 6 (VPC-6), ionization potential (IP), number of hydrogen acceptors (nHBAcc)	
Gajewicz et al., 2015	Enthalpy of formation of metal oxide nanocluster representing a fragment of the surface (ΔH_f^c), Mulliken's electronegativity of the nanocluster (χ^c)	Gajewicz et al., 2015
Pan et al., 2016	Molecular weight, cationic charge, mass percentage of metal elements, individual size, aggregation size	
Sizochenko et al., 2014	Unbonded two-atomic fragments [Me] · · ·[Me] (S_1), Wigner-Seitz radius of oxide's molecule (r_w), mass density (ρ), covalent index of the metal ion (CI), SiRMS-derived number of oxygen's atoms in a molecule (S_2), aggregation parameter (AP)	
Epa et al., 2012	Core material (I_{Fe3O4}), surface coating ($I_{dextran}$), surface charge ($I_{surfchg}$)	Shaw et al., 2008
Fourches et al., 2010; Liu et al., 2013a; Singh and Gupta, 2014	Size, R1 relaxivity, R2 relaxivity, zeta potential	

Table 3.6. (Continued)

Pathakoti et al., 2014	Absolute electronegativity of the metal atom (QMELECT), absolute electronegativity of the metal oxide (LZELEHHO), literature molar heat capacity of the metal oxide at 298.15 K (Cp), average of the alpha and beta LUMO energies of the metal oxide (ALZLUMO)	Pathakoti et al., 2014
Puzyn et al., 2011	Enthalpy of formation of a gaseous cation having the same oxidation state as that in the metal oxide structure (ΔH_{Me+})	Puzyn et al., 2011
Kar et al., 2014b	Charge of the metal cation corresponding to a given oxide (χ_{ox}), metal electronegativity (χ)	
Mu et al., 2016	Enthalpy of formation of a gaseous cation having the same oxidation state as that in the metal oxide structure (ΔH_{Me+}), polarization force (Z/r)	
Pan et al., 2016	Molecular weight, cationic charge, mass percentage of metal elements, individual size, aggregation size	
Singh and Gupta, 2014	Oxygen percent, molar refractivity, polar surface area	
Sizochenko et al., 2014	Unbonded two-atomic fragments [Me] · · ·[Me] (S_1), Wigner-Seitz radius of oxide's molecule (r_w), mass density (ρ), cation polarizing power (CPP), SiRMS-derived number of oxygen's atoms in a molecule (S_2), Tri-atomic fragments [Me]-[O]-[Me] (S_3), proportion of surface molecules to molecules in volume (SV)	
Zhang et al., 2012	Conduction band energy (E_c), solubility of metals	Zhang et al., 2012
Liu et al., 2013b	Ionic index of metal cation (Z^2/r), ENM conduction band energy (E_c), metal oxide ionization energy (ΔH_{IE}), metal oxide electronegativity (χ_{MeO}), atomization energy of metal oxide (E_{Atmz}), primary size (d), atomic mass of ENM metal (m_{Me})	
Sizochenko et al., 2015	Mass density, molecular weight, aligned electronegativity, covalent index, cation polarizing power, Wigner-Seitz radius, surface area, surface-area-to-volume ratio, aggregation parameter, two-atomic descriptor of van-der-Waals interactions, tri-atomic descriptor of atomic charges, tetra-atomic descriptor of atomic charges, size in DMEM	
Papa et al., 2015	Size of ENMs (X0), size in water (X1), size in phosphate buffered saline (X2), concentration (X4), zeta potential (X5)	Sayes and Ivanov, 2010
Liu et al., 2011	Size of ENM (d), volume concentration (θ_k), period of the ENM metal in the periodic table (P_{Me}), atomization energy of the metal oxide (E_{MeO})	Liu et al., 2011
Chen et al., 2016	Molecular polarizability, accessible surface area, solubility	Others
Luan et al., 2014; Kleandrova et al., 2014	Molar volume, polarizability, size of ENMs, electronegativity, hydrophobicity and polar surface area of surface coatings	

3.4 Interpret mechanisms of ENM biological activities with developed models

To enable the fast and inexpensive high-throughput prediction of diverse biological effects caused by ENMs, reliable nano-(Q)SARs should be based on mechanistic knowledge (OECD, 2007). Only when information on the underlying mechanisms is incorporated in modeling, proper and reliable extrapolation towards untested ENMs or organisms can be performed. Based on existing experimental data related to the cellular uptake of ENMs as well as the toxicity of ENMs to different cell lines and to *E. coli*, various nano-(Q)SARs were developed (Table 3.1). The significant descriptors introduced in the aforementioned nano-(Q)SAR studies are shown to be able to provide vital structural information on the factors affecting ENMs' cellular uptake and toxicity. Therefore, information on these descriptors as summarized in Table 3.6 is linked to the current understanding of the mechanisms of nanotoxicity.

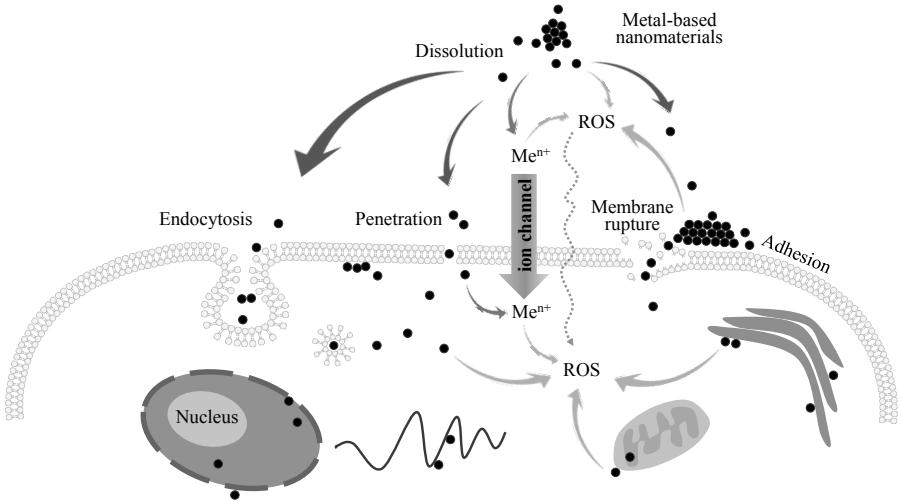


Figure 3.1. Overview of hypotheses associated with the responses of cellular membrane to the introduction of ENMs. It is assumed that endocytosis, penetration, adhesion of ENMs upon the cellular membrane, and cellular membrane rupture could possibly occur. Cellular membrane rupture is also considered to lead to the internalization of ENMs via the damage sites. Scenario of relevant ion release from ENMs, generation of reactive oxygen species (ROS), and ENMs-contacted interactions are also depicted.

3.4.1 Cellular uptake of ENMs

Once entering into the medium, ENMs may undergo various extra-and intracellular physical-chemical reactions such as dissolution, ion release, reactive oxygen species (ROS) generation, interaction with subcellular structures (e.g., cellular membrane, mitochondrion), and internalization into the cells (Figure 3.1). Cellular uptake of ENMs is always seen as an important process of ENMs' internalization and subsequently initiating the ENM contact-mediated or dissolved ion-associated intracellular reactions. As hypothesized, ENMs are conventionally transported into cells through endocytosis, a form of active transport in which cells take in ENMs by engulfing them (Zhao et al., 2011). Possible endocytotic processes proposed include phagocytosis, macropinocytosis, caveolae-dependent and clathrin-mediated endocytosis, and non-clathrin-, non-caveolae-mediated endocytosis (Unfried et al., 2007; Zhao et al., 2011). Besides, other responses of cellular membranes to adsorption of ENMs were also shown to exist. On the basis of a dissipative particle dynamics simulation study, Yue and Zhang (2011) concluded that surface adhesion, membrane penetration, and even ENM-induced membrane rupture could occur upon the ENM attachment to cellular membranes. Lin et al. (2010) and Xia et al. (2008) also demonstrated that ENMs could access to the cellular interior through direct membrane penetration.

In these internalization processes, surface properties of ENMs are essential for the ENM-biomolecule interactions and are deemed to be able to alter the cellular uptake pathways. In the experiment of Weissleder et al. (2005), a diversity of cellular uptake processes was observed especially for the PaCa2 cells. These authors consequently concluded that the translocation process is highly dependent on the surface modification of the ENMs. The studies showed that the lipophilicity of the surface molecules is an important discriminating factor that determines the chemical ability to interact with the lipid core of membranes (van de Waterbeemd et al., 2011). Fourches et al. (2010) reported that four descriptors (out of the top ten with the highest averaged frequency) SlogP_VSA0, SlogP_VSA1, SlogP_VSA2, and SlogP_VSA5 are intimately correlated with molecular lipophilicity of surface compounds. The ENMs with a higher PaCa2 cellular uptake are generally highly enriched for lipophilic surface modification (higher descriptor value). This is consistent with the results of Epa et al. (2012) in which C-005 (associated with hydrophobicity) was observed as a factor affecting ENMs' cellular uptake. Further confirmation was obtained by the appearance of *Atype* – N – 66 and *Atype* – N – 67 (Kar et al., 2014a) in a PLS model, and number of lipophilic groups (CH_2, fused rings) in the consensus model of Chau and Yap (2012). Also, the hydrogen bonding capacity of surface modifiers was explained to be one of the driving factors of ENMs' membrane penetrability (Chau and Yap, 2012). An ANN model predicting cellular uptake of ENMs by HUVEC was reported to include the

descriptors nRCONHR and nArOCON which characterize molecular hydrogen bonding capacity (Epa et al., 2012). In the same study, nN, nArOH, H-053, and O-058 were also found in the MLR model and were likewise interpreted as affecting the capability of H-bonding. In other nano-(Q)SARs, descriptors considered to correlate with this factor include nHDon (Ghorbanzadeh et al., 2012), WPSA-2, nHBDon, and nHBAcc (Singh and Gupta, 2014). Hence, these informative descriptors found in the developed nano-(Q)SARs confirmed again the previous experimental observations, that the lipophilicity of surface compounds is of significant importance for the cellular uptake of ENMs.

Additionally, shape, size, and flexibility of the surface compounds also play an important role in determining ENMs' passive transport across biological membranes. For instance, descriptors (not exclusively) characterizing molecular branching were constantly observed in the studies such as nR10, nCIR, nCs (Epa et al., 2012), *Wap* (Kar et al., 2014a), GATS1v, Mor29u, Mor14u (Ghorbanzadeh et al., 2012), SP-5, VP-4, and VPC-6 (Singh and Gupta, 2014). The Mor29u, Mor14u, SP-5, VP-4, and VPC-6 meanwhile also contain information of the molecular three-dimensional structures (e.g., mass, size, flexibility, and overall shape). Other relevant descriptors are namely ASP, DISPm, QZZm, QYYp, SPAM (Epa et al., 2012), $\sum\beta'$ (Kar et al., 2014a), De (Ghorbanzadeh et al., 2012), MOMI-XZ, nRotB (Singh and Gupta, 2014). Moreover, impacts on ENMs' cellular uptake were also reported to derive from the molecular reactive surface and electronegativity. Molecular reactive surface-related descriptors in the nano-(Q)SARs are vsa_don, vsa_other, vsa_base, PEOE_VSA_FPOS, PEOE_VSA+1 (Fourches et al., 2010), *Jurs–RPCS* (Kar et al., 2014a), WNSA-3 (Singh and Gupta, 2014). Descriptors associated with molecular electronegativity were observed to be nRNO2 (Kar et al., 2014a), primary, secondary, and tertiary N, halogens (Chau and Yap, 2012). It is not surprising that these factors of ENM surface modifiers may influence the ENM-biosurface interactions and pose effects on the cellular uptake of ENMs, independently or cooperatively. Either shape or size, or flexibility of surface modifiers of ENMs would affect the interactions between these molecules and the molecular sites of biosurfaces, change the conformation of binding complexes, and ultimately mediate the subsequent ENM-biosurface reactions in which the nature of the reactive surface and electronegativity also play a role.

As seen in Figure 3.1, internalization of ENMs into cells is generally considered as a crucial biological process triggering nanotoxicity. However, the adsorption of ENMs on cellular membranes may also affect cellular membrane integrity and lead to the formation of detects through the membranes (Lin et al., 2010; Thevenot et al., 2008; Xia et al., 2008; Yue and Zhang, 2011). This could probably result in the direct internalization of ENMs through the damage sites of membranes and the release of intracellular components that causes cell death. Notably, extracellular release of ions and formation of ROS are also considered to be

factors affecting the toxicity of ENMs in some cases. von Moos and Slaveykova (2014) reported that intracellular ROS generation can be stimulated by the presence of extracellular ROS as a response. The released ions and derived ROS may as well interact with cellular membranes, and dependently and/or independently influence ENMs' cellular uptake. This gives a possible explanation on the presence of the descriptor ionization potential (IP) in the nano-SAR of Singh and Gupta (2014), and may also be able to explain why the electronegativity-related descriptors nRNO2 (Kar et al., 2014a), primary, secondary, and tertiary N, and the halogens (Chau and Yap, 2012) generally appeared in relevant nano-(Q)SAR studies.

3.4.2 ENMs-induced biological effects

It is well-known that ENMs are capable of eliciting adverse biological effects by directly or indirectly triggering a series of physical-chemical reactions and ultimately causing cell damage. Reportedly, toxicity of ENMs could occur via a single mechanism or via combinations of the following mechanisms: (i) direct interactions with subcellular structures or biomolecules (e.g., membranes, mitochondria, proteins, DNA) which could lead to, for instance, mitochondrial damage, denaturation of proteins, formation of corona; (ii) release of chemical constituents from ENMs such as metal ions; (iii) surface property-based chemical reactivity of ENMs, e.g., photochemical, catalytic and redox properties; (iv) Trojan-horse type mechanisms, so called intruders in which ENMs act as vectors for transporting toxic chemicals (Figure 3.2).

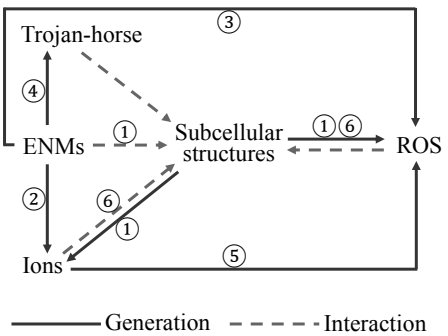


Figure 3.2. Schematic illustration of possible mechanisms of ENMs triggering nanotoxicity. 1) ENMs directly in contact with subcellular structures, which could promote the release of ions and ROS generation; 2) ENMs releasing ions; 3) ENMs contact-mediated ROS generation; 4) Trojan-horse mechanism triggered by ENMs; 5) Released ions increasing the formation of ROS; 6) Ion-dependent interactions which may lead to cellular damage or trigger ROS formation.

Generally, there is no doubt that metal-ions leaching from ENMs could act as a key factor causing biological effects of ENMs. Once the ENMs release dissolved ions surrounding the cells, it is often difficult to experimentally distinguish the effects caused by conventional metal ion release from the nano-specific effects. In such a context, the toxicity induced by ENMs is always considered to be intimately correlated with ENM dissolution. Comparable results on the toxicity of ZnO ENMs and Zn salts have been observed for the examples of *Pseudokirchneriella subcapitata* (Franklin et al., 2007), *Thamnocephalus platyurus* and *Daphnia magna* (Heinlaan et al., 2008; Wiench et al., 2009), and *E. coli* (Li et al., 2011). Result of studies on the toxicity of CuO ENMs to multiple species was also in agreement with this conclusion (Bondarenko et al., 2012; Heinlaan et al., 2008). It is commonly believed that ion release could occur after the cellular internalization of ENMs which consequently results in different mechanistic pathways of nanotoxicity. For instance, Stohs and Bagchi (1995) proposed a Haber-Weiss-Fenon cycle describing the stimulation of ion-leaching to ROS generation, taking Cu^{2+} as an example:

$$O_2^{\bullet -} + Cu^{2+} \rightarrow O_2 + Cu^{+}$$

$$Cu^{+} + H_2O_2 \rightarrow Cu^{2+} + OH^{-} + OH^{\bullet}$$

where the ROS such as superoxide anion radicals ($O_2^{\bullet -}$) could be derived from the one-electron reduction of molecular oxygen O_2:

$$O_2 + e \rightarrow O_2^{\bullet -}$$

In the Haber-Weiss-Fenton cycle, Cu^{2+} acts as catalysts of the formation of hydroxyl radicals which enhances the generation of ROS. Meanwhile, it was suggested that the release of ions could be accompanied by ROS formation as well such as in the Fenton reaction (Gajewicz et al., 2015):

$$Fe + O_2 + 2H^{+} \rightarrow Fe^{2+} + H_2O_2$$

$$Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + OH^{\bullet} + OH^{-}$$

Evidences from nano-(Q)SAR studies also demonstrated the contribution of ion release to nanotoxicity. Influence of metal solubility on nanotoxicity was indicted by the developed models (Zhang et al., 2012; Chen et al., 2016). Puzyn et al. (2011) developed a linear model based on the sole descriptor ΔH_{Me+} predicting toxicity of metal oxide ENMs to *E. coli*. It was explained that ΔH_{Me+} is an efficient descriptor characterizing the stability of metal oxides, which is associated with both the lattice energy of oxides and the sum of the

ionization potentials of a given metal. The release of cations with smaller charge is seen as more energetically favorable than that with larger charge (Mu et al., 2016). This explains the observations of previous studies giving an order of oxides toxicity as : $Me^{2+} > Me^{3+} > Me^{4+}$ (Puzyn et al., 2011). According to Kar et al. (2014b), the charge of the metal cation corresponding to a given oxide (χ_{ox}) was also used for the parameterization of nanotoxicity data. In the study of Liu et al. (2013b), the descriptor ionic index of metal cation Z^2/r was involved in building classification nano-SARs. Z is the ionic charge and r represents the Pauling ionic radius of the released ions (Pan et al., 2016). Z^2/r is a measure of the involvement of a metal ion into electrostatic interactions, and is able to provide information on the affinity of a metal ion for water molecules. Likewise, such form of index was also used in random forest models (Sizochenko et al., 2014; 2015), coupled with a parameter (S_1) describing the van der Waals interactions between surface molecules or cations. Other descriptors related to ionic charge and/or radius are polarization force (Mu et al., 2016), covalent index, tri-atomic descriptor of atomic charges, tetra-atomic descriptor of atomic charges (Sizochenko et al., 2015).

Accordingly, Gajewicz et al. (2015) employed two descriptors, i.e., enthalpy of formation of metal oxide nanocluster (ΔH_f^c) and the Mulliken's electronegativity of the cluster (χ^c), to linearly explain the cytotoxicity of metal oxide ENMs to HaCaT. The ΔH_f^c is associated with the energy of a single metal-oxygen bond in oxides ($E_{\Delta H}{}^o$) which can be expressed as (Portier et al., 2004):

$$E_{\Delta H^o} = \frac{2\Delta H_f^o \cdot 2.612 \times 10^{19}}{N_A n_e}$$

where N_A is the Avogadro number and n_e is the number of electrons involved in the formation reaction. A high $\Delta H_f{}^c$ value indicates a strongly bound cation of large formal charge in the oxides, and thus affects the detachment of metal cation from the surface of the ENMs. As for χ^c, Burello and Worth (2011) introduced that the electronegativity value of metal oxide (χ_{oxide}) can be calculated from that of the corresponding cation (χ_{cation}) using the equation (Portier et al., 2001):

$$\chi_{oxide} \approx 0.45\chi_{cation} + 3.36$$

Therefore, a higher value of χ_{cation} indicates a stronger ability of a cation to attract electrons in the Haber-Weiss-Fenton cycle which in turn results in higher reactivity of the metal oxide ENMs (Gajewicz et al., 2015). The two descriptors ΔH_f^c and χ^c meanwhile also refer to ENMs' surface redox activity. Burello and Worth (2011) reported that energy of a band gap (E_g) can be obtained based on the ΔH_f^0:

$$E_g = Ae^{0.34\Delta H_f^0}$$

and thus the conduction and valence band energies of oxides become:

$$E_c = -\chi_{oxide} + 0.5E_g + E_{shift}$$

$$E_v = -\chi_{oxide} - 0.5E_g + E_{shift}$$

where E_{shift} represent the value changes of band edges in respect to the solution's pH. As hypothesized, the redox potentials of relevant biological reactions could be unbalanced if they lie closer to the E_c or E_v, thereby causing cellular oxidative stress (Zhang et al., 2012). This was confirmed by Liu et al. (2013b) who identified E_c and χ_{oxide} for the development of nano-(Q)SAR models, and Kar et al. (2014b), Kleandrova et al. (2014), Luan et al. (2014), and Sizochenko et al. (2015) who used metal electronegativity as one of the modeling parameters. Pathakoti et al. (2014) as well obtained two descriptors (absolute electronegativity of the metal and metal oxide) for describing the toxicity of metal oxide ENMs to *E. coli* under darkness. Other descriptors considered to be associated with the surface redox properties of ENMs and causing oxidative stress are namely ΔH_{IE}, E_{Amz} (Liu et al., 2011; 2013b), *CI*, S_3 (Sizochenko et al., 2014), *Cp*, ALZLUMO (Pathakoti et al., 2014), and polarizability (Chen et al., 2016; Kleandrova et al., 2014; Luan et al., 2014) in relevant nano-(Q)SAR studies.

On the other hand, other than the general consensus taking ion release and ROS generation as driving factors in nanotoxicity, it is evident that other mechanisms of toxicity also play a vital role in certain cases. Xiao et al. (2015) reported that for both Cu and ZnO ENMs, the particles *per se*, rather than the dissolved ions, provided the major contribution to the toxicity to *Daphnia magna* (26% and 31%, respectively). Similarly, Hua et al. (2014) also revealed a dominant contribution of ZnO ENMs over the Zn ion tested for zebrafish embryos, for which the dissolution-driven mechanism of ENMs toxicity apparently does not apply. More precisely, it was shown that the shape of ENMs significantly affect ENMs' toxicity, as needle-shaped ZnO ENMs were proven to be the most toxic to *Phaeodactylum tricornutum* as compared to morphologically different ENMs with equal solubility and ion release (Peng et al., 2011). Observations of nanotoxicity affected by the shape of ENMs were also reported for ZnO nanospheres, nanosticks, and cuboidal submicron particles (Hua et al., 2014). Computational studies proved the involvement of surface property-related descriptors in nano-(Q)SAR modeling, such as the surface area and coating (Chen et al., 2016; Epa et al., 2012; Singh and Gupta, 2014; Sizochenko et al., 2015), hydrophobicity and polar surface area of surface molecules (Kleandrova et al., 2014; Luan et al., 2014), surface-area-to-volume ratio (Sizochenko et al., 2015), zeta potential (Papa et al., 2015), and

the Wigner-Seitz radius of oxide's molecule which describes the available fraction of molecules on the surface of ENMs (Sizochenko et al., 2014; 2015). The Wigner-Seitz radius also relates to molecular weight and density, and therefore as well the molecular volume which all have been indicated in the models (Kleandrova et al., 2014; Luan et al., 2014; Pan et al., 2016; Sizochenko et al., 2015). Descriptors relating to ENMs size (Kleandrova et al., 2014; Liu et al., 2011; 2013a,b; Luan et al., 2014; Pan et al., 2016; Papa et al., 2015; Singh and Gupta, 2014), material composition (Epa et al., 2012; Liu et al., 2011; Pan et al., 2016; Singh and Gupta, 2014; Sizochenko et al., 2014), and aggregation behaviors (Pan et al., 2016; Papa et al., 2015; Sizochenko et al., 2014; 2015) were also concluded to affect nanotoxicity (Suresh et al., 2013) from the aspects of relevant computational studies.

As mentioned above, ENMs may induce toxicity by direct steric hindrance or by binding with important reaction sites, or by indirect behaviors such as ion release, ENM surface-contacted interactions, or by acting as carriers for toxic chemicals (as in Figure 3.2). Take the case of TiO_2 as a typical example of ENM surface-mediated photochemical reaction, in which detachment of an electron could be activated by solar radiation (Kar et al., 2014b):

$$TiO_2 \overset{h\nu}{\Rightarrow} TiO_2^+ + \bar{e}$$

$$\bar{e} + O_2 \rightarrow O_2^{\bullet -}$$

$$O_2^{\bullet -} + 2H^+ + \bar{e} \rightarrow H_2O_2$$

$$O_2^{\bullet -} + H_2O_2 \rightarrow OH^{\bullet} + OH^- + O_2$$

$$H^+ + H_2O \Rightarrow OH^{\bullet} + H^+$$

The binding of ENMs with organelles could also cause a release of ions from interior storage due to the loss of membrane integrity. Unfried et al. (2007) reported that ENMs interacting with mitochondria are able to promote the release of interior-stored Ca^{2+}. The released ions are capable of triggering ROS production by direct catalysis, e.g., the Haber-Weiss-Fenton cycle, or indirect interference of biological functions such as interrupting the mitochondrial electron transduction (von Moos and Slaveykova, 2014). Besides, the ions *per se* could unbalance intracellular biological functions, eliciting inflammation, lysosomal damage, and inhibiting cellular respirations (He et al., 2014). The interactions of ENMs with subcellular structures (e.g., membrane-bound enzymes) were also shown to be capable of enhancing ROS production. Interestingly, presence of extracellular ROS was reported to be able to elevate intracellular ROS generation as depicted in Figure 3.2 (von Moos and Slaveykova, 2014).

In summary, the characteristics of ENMs may pose effects on the toxicity of ENMs as related to a single mechanism or to combinations of possible mechanisms. Analysis of the descriptors discussed in existing nano-(Q)SAR studies assists in offering statistical overview extracted from the complicated mechanistic pathways, and enables a mechanistic interpretation on the basis of the main driving factors. As discussed above, ENMs' surface properties are vital for their uptake by cells concerning the lipophilicity, hydrogen bonding capacity, electronegativity, shape, size, and flexibility of the surface modifiers. As for ENM-triggered toxicity, properties correlating with the ability of ion release and ROS generation could be important indicators, along with the information about the size, surface redox properties, and composition of ENMs.

3.5 Conclusions and Outlook

Enabling the development of reliable nano-(Q)SARs is capable of reducing the time and cost needed for conventional experimental evaluations, and thus benefits the risk evaluation and assessment of ENMs for regulatory purposes. Even though the promising potential of extending (Q)SARs into nanotoxicity has been addressed, the nano-(Q)SAR approach is still in its infancy. As far as it is understood, scarcity of (properly documented) experimental data is regarded as one of the major drawbacks in building nano-(Q)SAR models. The information provided in Table 3.2 indicated a very limited availability of existing data as only a few datasets constantly appeared in the overview of nano-(Q)SAR studies, in spite of the numerous scientific programs on ENMs' safety and design. This suggests that (i) most of the studies reported do not meet the modeling criteria which, amongst others include lack of relevant pristine or characterization data, lack of a description of the method used, or lack of reporting of a consistent endpoint; or (ii) the integration of existing experimental data based on various studies is currently lacking, which hinders the inclusion of this valuable information into the nano-modeling field (Chen et al., 2017). Therefore, in light of advancing computational nanotoxicology, a summary and also organization of potentially useful nanotoxicity is essential. Besides the data quantity, the quality of experiment data collected should also be taken care of for the data-driven nano-(Q)SAR approach, which was found absent in the relevant studies owing to the single-source strategy of retrieving data for a model. It is suggested that the quality of experimental information assembled from various sources ought to be evaluated by suitable tools before model construction. This is seen as helpful for improving the statistical significance and predictability of a model.

Meanwhile, the grouping and characterization of ENMs as well remain crucial for developing nano-(Q)SARs. In general, the strategies of grouping ENMs are considered to

be ENM composition-based, toxic mode of action-based, or further clustering-based (Fourches et al., 2010; Liu et al., 2013a). Characterization of ENMs will subsequently be carried out for the ENM groups in terms of molecular structural descriptors. However, concerns have always been expressed with regard to the question whether it is possible to build nano-(Q)SARs without considering ENMs' dynamic transformations in the exposure medium. On the one hand it is well-known that once entering into a medium, ENMs are more likely to strongly react with the components of the test medium and undergo dramatic changes of surface properties. These surface transformations would in return affect ENMs' reactivity and subsequent biological behaviors (e.g., cellular uptake, interaction with subcellular structures). In such a context, modeling based solely on the information of ENMs' pristine structures could be biased and could result in poor predictability and reliability of the models generated. Meanwhile, on the other hand, a few efforts did provide evidences regarding the feasibility of building nano-models using the characteristics of pristine ENMs (Kar et al., 2014a; Pathakoti et al., 2014; Puzyn et al., 2011; Singh and Gupta, 2014; Sizochenko et al., 2014). Actually the possibility exists that the characteristics of pristine ENMs influence the biological effects of ENMs by affecting ENMs' dynamic transformations in media, and it can be hypothesized that even though changes of ENM property could occur in the exposure media, the characteristics of the pristine ENMs still are linked to adverse biological effects of ENMs. In this circumstance, constructing nano-(Q)SARs with only characteristics of pristine ENMs could enable the development of high-throughput protocols for non-testing nanotoxicity evaluation. This is expected to allow to reduce the high cost and time needed by conventional evaluation methods. However, all the proposed hypotheses should be further confirmed by more nano-(Q)SAR studies in pace with the advance of computational nanotoxicology.

In conclusion, the added value of this review can be summarized as:

(i) a general overview was provided of the datasets being widely used in nano-(Q)SAR studies coupled with the provided characterization of ENMs. Experimental data were shown to be mainly available concerning the cellular uptake by different cell lines (e.g., PaCa2, HUVEC), cytotoxicity to cells (e.g., HaCaT, BEAS-2B), and the toxicity to *E. coli*. A limited usage of existing data in relevant investigations was observed;

(ii) an overview was presented on nano-(Q)SARs developed so far, based on a variety of modeling techniques such as linear and non-linear regressions (MLR, PLS, logistic regression), SVM, ANN, and *k*NN;

(iii) an interpretation of the underlying mechanisms of ENMs' toxicity and cellular uptake was provided on the basis of the descriptors discussed in nano-(Q)SAR studies. Surface

properties of ENMs were deemed vital for the uptake of ENMs by different cell lines, such as lipophilicity, hydrogen bonding capacity, electronegativity, shape, and size. The capability of releasing ions and generating ROS, surface redox properties of ENMs were concluded to be important indicators for evaluating the toxicity of ENMs;

(iv) an outlook was presented regarding the experimental data needed for future modeling and the need of proper characterization of ENMs. Owing to the limited data availability, optimizing the usage of existing information of nanotoxicity should be deliberately considered, and thus integrating relevant available data becomes vital for the development of nano-(Q)SARs. Meanwhile, whether or not the dynamic transformations of ENMs in media play a vital role in the computer-aided nanotoxicity also ought to be further discussed.

Acknowledgements

The Chinese Scholarship Council (CSC) is gratefully acknowledged for its financial support to Guangchao Chen and Yinlong Xiao. The research described in this work was supported by the European Union's European Union Seventh Framework Programme under EC-GA No. 604602 'FUTURENANONEEDS'. Martina Vijver was funded by NWO VIDI 864.13.010.

Reference

Asare N, Instanes C, Sandberg WJ, Refsnes M, Schwarze P, Kruszewski M, Brunborg G. Cytotoxic and genotoxic effects of silver nanoparticles in testicular cells. Toxicology. 2012, 291:65-72.

Bondarenko O, Ivask A, Kakinen A, Kahru A. Sub-toxic effects of CuO nanoparticles on bacteria: kinetics, role of Cu ions and possible mechanisms of action. Environ Pollut. 2012, 169:81-9.

Burello E, Worth AP. A theoretical framework for predicting the oxidative stress potential of oxide nanoparticles. Nanotoxicology. 2011, 5:228-35.

Chau YT, Yap CW. Quantitative Nanostructure–Activity Relationship modelling of nanoparticles. RSC Adv. 2012, 2:8489-8496.

Chen G, Peijnenburg WJ, Kovalishyn V, Vijver MG. Development of nanostructure–activity relationships assisting the nanomaterial hazard categorization for risk assessment and regulatory decision-making. RSC Adv. 2016, 6:52227-52235.

Chen G, Peijnenburg WJ, Xiao Y, Vijver MG. Developing species sensitivity distributions for metallic nanomaterials considering the characteristics of nanomaterials, experimental conditions, and different types of endpoints. Food Chem Toxicol. 2017. doi: 10.1016/j.fct.2017.04.003.

Chen G, Vijver MG, Peijnenburg WJ. Summary and analysis of the currently existing literature data on metal-based nanoparticles published for selected aquatic organisms: Applicability for toxicity prediction by (Q)SARs. Altern Lab Anim. 2015, 43:221-40.

Ellegaard-Jensen L, Jensen KA, Johansen A. Nano-silver induces dose–response effects on the nematode *Caenorhabditis elegans*. Ecotoxicol Environ Saf. 2012, 80:216-23.

Epa VC, Burden FR, Tassa C, Weissleder R, Shaw S, Winkler DA. Modeling Biological Activities of Nanoparticles. Nano Lett. 2012, 12:5808-12.

Fernández A, Lombardo A, Rallo R, Roncaglioni A, Giralt F, Benfenati E. Quantitative consensus of bioaccumulation models for integrated testing strategies. Environ Int. 2012, 45:51-8.

Fourches D, Pu D, Tassa C, Weissleder R, Shaw SY, Mumper RJ, Tropsha A. Quantitative Nanostructure-Activity Relationship Modeling. ACS Nano. 2010, 4:5703-12.

Fourches D, Pu D, Tropsha A. Exploring Quantitative Nanostructure-Activity Relationships (QNAR) Modeling as a Tool for Predicting Biological Effects of Manufactured Nanoparticles. Comb Chem High Throughput Screen. 2011, 14:217-25.

Franklin NM, Rogers NJ, Apte SC, Batley GE, Gadd GE, Casey PS. Comparative toxicity of nanoparticulate ZnO, bulk ZnO, and $ZnCl_2$ to a freshwater microalga (*Pseudokirchneriella subcapitata*): the importance of particle solubility. Environ Sci Technol. 2007, 41:8484-90.

Gajewicz A, Rasulev B, Dinadayalane TC, Urbaszek P, Puzyn T, Leszczynska D, Leszczynski J. Advancing risk assessment of engineered nanomaterials: application of computational approaches. Adv Drug Deliv Rev. 2012, 64:1663-93.

Gajewicz, A. Schaeublin, N. Rasulev, B. Hussain, S. Leszczynska, D. Puzyn, T. Leszczynski, J. Towards understanding mechanisms governing cytotoxicity of metal oxides nanoparticles: Hints from nano-QSAR studies. Nanotoxicology. 2015, 9:313-25.

Ghorbanzadeh M, Fatemi MH, Karimpour M. Modeling the Cellular Uptake of Magnetofluorescent Nanoparticles in Pancreatic Cancer Cells: A Quantitative Structure Activity Relationship Study. Ind Eng Chem Res. 2012, 51:10712-10718.

Gramatica P. Principles of QSAR models validation: internal and external. QSAR Comb Sci. 2007, 26:694-701.

He X, Aker WG, Leszczynski J, Hwang HM. Using a holistic approach to assess the impact of engineered nanomaterials inducing toxicity in aquatic systems. J Food Drug Anal. 2014, 22:128-46.

Heinlaan M, Ivask A, Blinova I, Dubourguier H-C, Kahru A. Toxicity of nanosized and bulk ZnO, CuO and TiO_2 to bacteria *Vibrio fischeri* and crustaceans Daphnia magna and Thamnocephalus platyurus. Chemosphere. 2008, 71:1308-16.

Hu X, Cook S, Wang P, Hwang HM. In vitro evaluation of cytotoxicity of engineered metal oxide nanoparticles. Sci Total Environ. 2009, 407:3070-2.

Hua J, Vijver MG, Richardson MK, Ahmad F, Peijnenburg WJ. Particle-specific toxic effects of differently shaped zinc oxide nanoparticles to zebrafish embryos (*Danio rerio*). Environ Toxicol Chem. 2014, 33:2859-2868.

Ivask A, Juganson K, Bondarenko O, Mortimer M, Aruoja V, Kasemets K, Blinova I, Heinlaan M, Slaveykova V, Kahru A. Mechanisms of toxic action of Ag, ZnO and CuO nanoparticles to selected ecotoxicological test organisms and mammalian cells in vitro: a comparative review. Nanotoxicology. 2014, 8:57-71.

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 In Vitro. 2014a, 28:600-6.

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. 2014b, 107:162-9.

Kar S, Roy K. QSAR modeling of toxicity of diverse organic chemicals to *Daphnia magna* using 2D and 3D descriptors. J Hazard Mater. 2010, 177:344-51.

Kar S, Roy K. First report on development of quantitative interspecies structure–carcinogenicity relationship models and exploring discriminatory features for rodent carcinogenicity of diverse organic chemicals using OECD guidelines. Chemosphere. 2012, 87:339-55.

Kleandrova VV, Luan F, González-Díaz H, Ruso JM, Speck-Planche A, Cordeiro MN. 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. 2014, 48:14686-94.

Kumari M, Khan SS, Pakrashi S, Mukherjee A, Chandrasekaran N. Cytogenetic and genotoxic effects of zinc oxide nanoparticles on root cells of *Allium cepa*. J Hazard Mater. 2011, 190:613-21.

Li M, Zhu L, Lin D. Toxicity of ZnO nanoparticles to *Escherichia coli*: mechanism and the influence of medium components. Environ Sci Technol. 2011, 45:1977-83.

Lin J, Zhang H, Chen Z, Zheng Y. Penetration of Lipid Membranes by Gold Nanoparticles: Insights into Cellular Uptake, Cytotoxicity, and Their Relationship. ACS Nano. 2010, 4:5421-9.

Linkov I, Steevens J, Adlakha-Hutcheon G, Bennett E, Chappell M, Colvin V, Davis JM, Davis T, Elder A, Foss Hansen S, Hakkinen PB, Hussain SM, Karkan D, Korenstein R, Lynch I, Metcalfe C, Ramadan AB, Satterstrom FK. Emerging methods and tools for environmental risk assessment, decision-making, and policy for nanomaterials: summary of NATO Advanced Research Workshop. J Nanopart Res. 2009, 11:513-527.

Liu R, Rallo R, George S, Ji Z, Nair S, Nel AE, Cohen Y. Classification NanoSAR development for cytotoxicity of metal oxide nanoparticles. Small. 2011, 7:1118-1126.

Liu R, Rallo R, Weissleder R, Tassa C, Shaw S, Cohen Y. Nano-SAR development for bioactivity of nanoparticles with considerations of decision boundaries. Small. 2013a, 9:1842-52.

Liu R, Zhang HY, Ji ZX, Rallo R, Xia T, Chang CH, Nel A, Cohen Y. Development of structure-activity relationship for metal oxide nanoparticles. Nanoscale. 2013b, 5:5644-53.

Long TC, Saleh N, Tilton RD, Lowry GV, Veronesi B. Titanium dioxide (P25) produces reactive oxygen species in immortalized brain microglia (BV2): implications for nanoparticle neurotoxicity. Environ Sci Technol. 2006, 40:4346-52.

Luan F, Kleandrova VV, González-Díaz H, Ruso JM, Melo A, Speck-Planche A, Cordeiro MN. Computer-aided nanotoxicology: assessing cytotoxicity of nanoparticles under diverse experimental conditions by using a novel QSTR-perturbation approach. Nanoscale. 2014, 6:10623-30.

Mu Y, Wu F, Zhao Q, Ji R, Qie Y, Zhou Y, Hu Y, Pang C, Hristozov D, Giesy JP, Xing B. Predicting toxic potencies of metal oxide nanoparticles by means of nano-QSARs. Nanotoxicology. 2016, 10:1207-14.

Netzeva TI, Worth A, Aldenberg T, Benigni R, Cronin MT, Gramatica P, Jaworska JS, Kahn S, Klopman G, Marchant CA, Myatt G, Nikolova-Jeliazkova N, Patlewicz GY, Perkins R, Roberts D, Schultz T, Stanton DW, van de Sandt JJ, Tong W, Veith G, Yang C. Current status of methods for defining the applicability domain of (quantitative) structure–activity relationships. Altern Lab Anim. 2005, 33:155-173.

Nirmala R, Park HM, Kalpana D, Kang HS, Navamathavan R, Lee YS, Kim HY. Bactericidal activity and in vitro cytotoxicity assessment of hydroxyapatite containing gold nanoparticles. J Biomed Nanotechnol. 2011, 7:342-50.

OECD. Guidance document on the validation of (quantitative) structure activity relationship [(Q)SAR] models. OECD Series on Testing and Assessment No. 69. ENV/JM/MONO(2007)2. Organization for Economic Cooperation and Development, Paris, France. 2007.

Pan Y, Li T, Cheng J, Telesca D, Zink JI, Jiang J. Nano-QSAR modeling for predicting the cytotoxicity of metal oxide nanoparticles using novel descriptors. RSC Adv. 2016, 6:25766-25775.

Papa E, Doucet JP, Doucet-Panaye A. Linear and non-linear modelling of the cytotoxicity of TiO2 and ZnO nanoparticles by empirical descriptors. SAR QSAR Environ Res. 2015, 26:647-65.

Pathakoti K, Huang MJ, Watts JD, He X, Hwang HM. 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. 2014, 130:234-40.

Peijnenburg W. Structure-activity relationships for biodegradation: A critical review. Pure Appl Chem. 2009, 66:1931-1941.

Peng X, Palma S, Fisher NS, Wong SS. Effect of morphology of ZnO nanostructures on their toxicity to marine algae. Aquat Toxicol. 2011, 102:186-96.

Portier J, Campet G, Poquet A, Marcel C, Subramanian MA. Degenerate semiconductors in the light of electronegativity and chemical hardness. Int J Inorg Mater. 2001, 3:1039-1043.

Portier J, Hilal HS, Saadeddin I, Hwang SJ, Subramanian MA, Campet G. Thermodynamic correlations and band gap calculations in metal oxides. Prog Solid State Chem. 2004, 32:207-217.

Puzyn T, Leszczynska D, Leszczynski J. Toward the Development of "Nano-QSARs": Advances and Challenges. Small. 2009, 5:2494-509.

Puzyn T, Rasulev B, Gajewicz A, Hu X, Dasari TP, Michalkova A, Hwang HM, Toropov A, Leszczynska D, Leszczynski J. Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles. Nat Nanotechnol. 2011, 6:175-8.

Raymond JW, Rogers TN, Shonnard DR, Kline AA. A review of structure-based biodegradation estimation methods. J Hazard Mater. 2001, 84:189-215.

Russell WMS, Burch RL. The Principles of Humane Experimental Technique. London: Methuen & Co. Ltd. 1959. [Reissued: 1992, Universities Federation for Animal Welfare Herts, England.]

Sayes C, Ivanov I. Comparative study of predictive computational models for nanoparticle-induced cytotoxicity. Risk Anal. 2010, 30:1723-34.

Sharma V, Singh SK, Anderson D, Tobin DJ, Dhawan A. Zinc oxide nanoparticle induced genotoxicity in primary human epidermal keratinocytes. J Nanosci Nanotechnol. 2011, 11:3782-8.

Shaw SY, Westly EC, Pittet MJ, Subramanian A, Schreiber SL, Weissleder R. Perturbational profiling of nanomaterial biologic activity. Proc Natl Acad Sci U S A. 2008, 105:7387-92.

Singh KP, Gupta S. Nano-QSAR modeling for predicting biological activity of diverse nanomaterials. RSC Adv. 2014, 4:13215-13230.

Sizochenko N, Rasulev B, Gajewicz A, Kuz'min V, Puzyn T, Leszczynski J. From basic physics to mechanisms of toxicity: the "liquid drop" approach applied to develop predictive classification models for toxicity of metal oxide nanoparticles. Nanoscale. 2014, 6:13986-93.

Sizochenko N, Rasulev B, Gajewicz A, Mokshyna E, Kuz'min VE, Leszczynski J, Puzyn T. Causal inference methods to assist in mechanistic interpretation of classification nano-SAR models. RSC Adv. 2015, 5:77739-77745.

Stewart JJP. Optimization of parameters for semiempirical methods. V. Modification of NDDO approximations and application to 70 elements. J Mol Model. 2007, 13:1173-213.

Stewart JJP. MOPAC2009. S.C. Chemistry, Editor 2009. Available at: http://openmopac.net/MOPAC2009.html.

Stohs SJ, Bagchi D. Oxidative mechanisms in the toxicity of metal ions. Free Radic Biol Med. 1995, 18:321-36.

Suresh AK, Pelletier DA, Doktycz MJ. Relating nanomaterial properties and microbial toxicity. Nanoscale. 2013, 5:463-74.

Sushko I, Novotarskyi S, Körner R, Pandey AK, Rupp M, Teetz W, Brandmaier S, Abdelaziz A, Prokopenko VV, Tanchuk VY, Todeschini R, Varnek A, Marcou G, Ertl P, Potemkin V, Grishina M, Gasteiger J, Schwab C, Baskin II, Palyulin VA, Radchenko EV, Welsh WJ, Kholodovych V, Chekmarev D, Cherkasov A, Aires-de-Sousa J, Zhang QY, Bender A, Nigsch F, Patiny L, Williams A, Tkachenko V, Tetko IV. Online chemical modeling environment (OCHEM): web platform for data storage, model development and publishing of chemical information. J Comput Aided Mol Des. 2011, 25:533-54.

Thevenot P, Cho J, Wavhal D, Timmons RB, Tang L. Surface chemistry influences cancer killing effect of TiO_2 nanoparticles. Nanomedicine. 2008, 4:226-36.

Thill A, Zeyons O, Spalla O, Chauvat F, Rose J, Auffan M, Flank AM. Cytotoxicity of CeO_2 nanoparticles for *Escherichia coli*. Physico-chemical insight of the cytotoxicity mechanism. Environ Sci Technol. 2006, 40:6151-6.

Toropov AA, Toropova AP, Benfenati E, Gini G, Puzyn T, Leszczynska D, Leszczynski J. Novel application of the CORAL software to model cytotoxicity of metal oxide nanoparticles to bacteria *Escherichia coli*. Chemosphere. 2012, 89:1098-102.

Toropov AA, Toropova AP, Puzyn T, Benfenati E, Gini G, Leszczynska D, Leszczynski J. QSAR as a random event: Modeling of nanoparticles uptake in PaCa2 cancer cells. Chemosphere. 2013, 92:31-37.

Tran N, Mir A, Mallik D, Sinha A, Nayar S, Webster TJ. Bactericidal effect of iron oxide nanoparticles on Staphylococcus aureus. Int J Nanomedicine. 2010, 5:277-83.

Unfried K, Albrecht C, Klotz L, von Mikecz A, Grether-Beck S, Schins R. Cellular responses to nanoparticles: Target structures and mechanisms. Nanotoxicology. 2007, 1:52-71.

van de Waterbeemd H, Smith DA, Jones BC. Lipophilicity in PK design: methyl, ethyl, futile. J Comput Aided Mol Des. 2001, 15:273-86.

von Moos N, Slaveykova VI. Oxidative stress induced by inorganic nanoparticles in bacteria and aquatic microalgae-state of the art and knowledge gaps. Nanotoxicology. 2014, 8:605-30.

Weissleder R, Kelly K, Sun EY, Shtatland T, Josephson L. Cell-specific targeting of nanoparticles by multivalent attachment of small molecules. Nat Biotechnol. 2005, 23:1418-23.

Wiench K, Wohlleben W, Hisgen V, Radke K, Salinas E, Zok S, Landsiedel R. Acute and chronic effects of nano- and non-nano-scale TiO_2 and ZnO particles on mobility and reproduction of the freshwater invertebrate Daphnia magna. Chemosphere. 2009, 76:1356-65.

Wiesner MR, Lowry GV, Alvarez P, Dionysiou D, Biswas P. Assessing the risks of manufactured nanomaterials. Environ Sci Technol. 2006, 40:4336-45.

Winkler DA. Recent advances, and unresolved issues, in the application of computational modelling to the prediction of the biological effects of nanomaterials. Toxicol Appl Pharmacol. 2016, 299:96-100.

Win-Shwe TT, Fujimaki H. Nanoparticles and neurotoxicity. Int J Mol Sci. 2011, 12:6267-6280.

Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics. Chemometr Intell Lab Syst. 2001, 58:109-130.

Worth AP. The role of QSAR methodology in the regulatory assessment of chemicals. In: Puzyn T, Leszczynski J, Cronin MTD, editors. Recent Advances in QSAR Studies: Methods and Applications. Springer; Heidelberg: 2010. p. 367-82.

Wu J, Wang C, Sun J, Xue Y. Neurotoxicity of silica nanoparticles: brain localization and dopaminergic neurons damage pathways. ACS Nano. 2011, 5:4476-89.

Xia T, Rome L, Nel A. Nanobiology: Particles slip cell security. Nat Mater. 2008, 7:519-20.

Xiao Y, Vijver MG, Chen G, Peijnenburg WJGM. Toxicity and Accumulation of Cu and ZnO nanoparticles in *Daphnia magna*. Environ Sci Technol. 2015, 49:4657-4664.

Yue T, Zhang X. Molecular understanding of receptor-mediated membrane responses to ligand-coated nanoparticles. Soft Matter. 2011, 7:9104-9112.

Zhang H, Ji Z, Xia T, Meng H, Low-Kam C, Liu R, Pokhrel S, Lin S, Wang X, Liao YP, Wang M, Li L, Rallo R, Damoiseaux R, Telesca D, Mädler L, Cohen Y, Zink JI, Nel AE. Use of metal oxide nanoparticle band gap to develop a predictive paradigm for oxidative stress and acute pulmonary inflammation. ACS Nano. 2012, 6:4349-68.

Zhao F, Zhao Y, Liu Y, Chang X, Chen C, Zhao Y. Cellular uptake, intracellular trafficking, and cytotoxicity of nanomaterials. Small. 2011, 7:1322-37.

CHAPTER 4

DEVELOPMENT OF NANOSTRUCTURE-ACTIVITY RELATIONSHIPS ASSISTING THE NANOMATERIAL HAZARD CATEGORIZATION FOR RISK ASSESSMENT AND REGULATORY DECISION-MAKING

Chen G, Peijnenburg WJGM, Kovalishyn V, Vijver MG

Published in *RSC Advances*. 2016, 6(57): 52227-52235

Abstract

Categorization of the environmental hazards associated with engineered nanomaterials (ENMs) is important for evaluating the potential risks brought by commercialized ENMs. Such a task is so far severely hindered because of insufficient amount of available toxicity data. As biological assays are costly and time-consuming and also face the ethical issue of animal use, computational modeling such as (quantitative) nanostructure-activity relationships (nano-(Q)SARs) is valued as a potential tool to fill in the data gaps. With this in mind, nano-SARs classifying the ecotoxicity of ENMs were developed in this study with the aims: (i) to examine the availability of nanoecotoxicity data in developing nano-SARs; and (ii) to build nano-SARs that assist the hazard categorization of ENMs for the regulatory purposes. Multi-source ecotoxicity data were retrieved, on basis of which descriptors quantifying the ENM structures were calculated. By employing four extensively used tree algorithms, global nano-SARs across species and species-specific models were derived with significant predictive power. For the LC50 global models, the functional tree, C4.5 decision tree, and random tree models all correctly classified more than 70.0% of the samples on training (320 ENMs) and test sets (80 ENMs). The functional tree predicting the toxicity of metallic ENMs to *Danio rerio* showed accuracies of 93.4% and 100% on respectively training (76 ENMs) and test sets (18 ENMs). Descriptors present in the species-specific models were analyzed to discuss the key factors affecting nanotoxicity. With easily obtained descriptors and transparent predictive rules, we believe the developed nano-SARs could assist the expedited review of ENMs' hazards and facilitate better-informed regulatory decisions of ENMs.

Key words: classification, mechanisms, metallic nanomaterials, nano-(Q)SAR, toxicity

4.1 Introduction

Assessing the potential environmental risks posed by engineered nanomaterials (ENMs) is essential to ensure that the marketed ENMs are used as safely as possible. It is believed that, a preliminary categorization of ENMs will benefit the early stages of qualitative risk analysis either by manufacturers or by regulators, to target the ENMs of high risk concerns and so as to prioritize more detailed testing of ENMs (Godwin et al., 2015). The European Chemicals Agency, for instance, has released reports and documents alike to address the usefulness of ENM grouping serving to the streamline testing for the regulatory purposes (Godwin et al., 2015). The U.S.-Canada Regulatory Cooperation Council also reported development of the classification scheme for ENMs in order to identify the 'ENMs of concern' that are likely to behave differently compared to their bulk scale counterparts (RCC-NI, 2013). Generally, one of the commonly used strategies of ENM categorization is to group ENMs based on different measures of biological activities. An example of this can be found in the CLP-Regulation (EC) No 1272/2008, which suggests that chemicals can be classified as acutely toxic or as chronically toxic at multiple levels according to the outcomes of standardized toxicity tests (Juganson et al., 2015; CEC, 1996). Another example is the EU Directive 93/67/EEC that recommends to rank the chemical hazard to aquatic species into four hierarchies, i.e. very toxic, toxic, harmful, and not classified, on the basis of at least three standard test species algae, crustacean, and fish (CEC, 1996). Unsurprisingly, however, those risk potential-based material categorizations require an enormous amount of hazard information of ENMs intended for adequately evaluating the safety of the materials. Given the substantial number of existing, non-tested ENMs and the rapid growth of ENMs innovation, it is, consequently, expected that alternatives of testing assays such as (quantitative) nanostructure-activity relationships (termed as nano-(Q)SARs) could be effectively used to fill in data gaps while with the minimum of financial cost and time consumption. The application of (Q)SARs in ENM categorization is seen to be quite advanced as it is capable of promoting the safe use of ENMs (Tantra et al., 2015). Meanwhile, employing (Q)SARs in ENMs' risk assessment also meets the 3R's principle (refine, reduce, and replace) of animal use in toxicity testing (Russell and Burch, 1959).

Previously, a few nano-(Q)SAR models have been established by linking ENMs' biological responses to the experimental and/or computational characterization of ENMs (Chen et al., 2015). One of the issues so far in developing nano-(Q)SARs is that a relatively small number of datasets were repeatedly used by different studies (Winkler, 2016). This may be because of one of the obstacles of using multi-source data in developing nano-(Q)SARs being the lack of data consistency between diverse researches. This lack of data leads to the difficulty of comprehensively characterizing the structures of ENMs in an entire dataset especially for fully quantifying the information on surface coatings and functional groups of

ENMs. However, given the constantly increasing amount of scientific resources from numerous scientific programs on nanomaterial safety, and given the urgent need of further development in computational nanotoxicology to assist the risk assessment of nanomaterials, nano-(Q)SARs based on the integration and maximization of the use of existing nanotoxicity data also seems to be of particular importance. We hence aimed to derive classification nano-SARs by using the currently available and accessible nanotoxicity data on environmental species shared in various publications and scientific resources. Feasible strategy of computationally characterizing the structures of ENMs was chosen. The purposes of this study are summarized as, firstly, to examine the availability of existing nanoecotoxicity data in developing nano-SARs; and secondly, to build classification models for ENMs assisting the nanomaterial hazard categorization for estimating the risks of metal-based nanomaterials.

To begin with, three datasets were obtained from various publications and scientific resources, and considered for the use of modeling. The structural descriptors were calculated using a web-based platform Online Chemical Modeling Environment (OCHEM) which characterize the information of the core of metal-based ENMs (Sushko et al., 2011). To acquire transparent and easily applicable classification models, four extensively employed tree algorithms embedded in the Weka (version 3.6) were considered for modeling, namely functional tree, C4.5 decision tree, random tree, and simple CART (Hall et al., 2009). Based on the descriptors and algorithms, global nano-SARs across species as well as species-specific models were developed with significant predictability. The global models are favorable for ranking the general biological effects of ENMs regardless of targeted species, while species-specific models are able to offer in depth knowledge of nanotoxicity and may also be more applicable when the estimation of nanotoxicity is based on certain species (e.g. categorize ENMs based on EU Directive 93/67/EEC). Descriptors appearing in the species-specific nano-SARs were analyzed in light of a mechanistic interpretation of the toxicity triggered by metallic ENMs. The present study examined the availability of published nanoecotoxicity data in deriving nano-(Q)SARs and demonstrated the possibility of building nano-SARs using multi-sources datasets.

4.2 Methods

4.2.1 Datasets

We previously established a database summarizing and describing the toxicity of metal-based ENMs to selected organisms in light of the development of nano-(Q)SARs (Chen et

al., 2015). Records of the commonly used toxicity endpoints in this database, including EC50 (the effective concentration that causes 50% response), EC20, LC50 (the concentration which leads to 50% mortality), LC20, MIC (minimum inhibitory concentration), and NOEC (no observed effect concentration) were manually uploaded to the web-based platform on 18th August, 2015 (Sushko et al., 2011). Using the OCHEM platform, an analysis of the available ecotoxicity data of metal-based ENMs was performed on 28th August, 2015, which provided us with three datasets containing the toxicity of various metal-based ENMs to different hierarchies of species: (I) 400 ENMs from 90 publications or reports provided with experimental data on LC50; (II) 450 ENM records from 79 publications or reports with quantitative information on EC50 values; and (III) 166 ENMs obtained from 13 publications with experimental values of the MIC. MIC characterizes the antimicrobial properties of ENMs and is therefore a common experimental endpoint in antimicrobial assays. Even though the use of MIC does not currently fit into the scheme of evaluating ENMs' risks based on different species, we still included this case study so as to further examine the feasibility of building nano-SARs for different hierarchies of species. Units of the toxicity values were unified into mg/L in the datasets. For building global nano-SARs across species, the three datasets I, II, and III were used as three case studies. As for constructing models for single species, from each of the dataset two species with the most toxicity endpoint records were chosen. As a result, the selected species were *Danio rerio* (94 records including embryo, LC50), *Daphnia magna* (102, LC50), *Pseudokirchneriella subcapitata* (66, EC50), *Daphnia magna* (105, EC50), *Escherichia coli* (41, MIC), and *Staphylococcus aureus* (39, MIC).

As it is acknowledged, thresholds that discretize the numeric values are of significant importance for building classification models, which thus should be carefully discussed and selected on the basis of different strategies and application requirements (Liu et al., 2013). In this study, we initially examined the tendency of model predictability with the shift of threshold values. And afterwards thresholds that lead to the most balanced predictive performances were conditionally considered. Referring to the regulations and directives nowadays in force, consideration of the thresholds for global models was restricted to the values of 0.1, 1.0, 10.0, and 100.0 mg/L, which are, for instance, used by both the aforementioned CLP-Regulation (EC) No 1272/2008 and the EU Directive 93/67/EEC. For the species specific nano-SARs, thresholds of 1.0, 10, 100 mg/L were taken (for *Escherichia coli* and *Staphylococcus aureus* only 10 and 100 mg/L because of narrower variation of toxicity values). Within each dataset the records were ranked based on the values of the toxicity endpoints. ENMs with toxicity values less than pre-specified threshold value were assigned to the 'active' class, and the rest of ENMs were labeled as 'inactive'. When building models, 20% of the dataset was exclusively utilized for external validation.

4.2.2 Descriptor calculation

Obtaining the structural descriptors of ENMs is essential to characterize the structures of ENMs besides the experimental measures. Using the 'Calculate descriptors' function implemented in OCHEM, three types of descriptors were calculated and acquired, the E-state, ALogPS, and Chemaxon descriptors. For the E-state, both atom and bond types were considered for the indices and counts descriptors during calculation. The selected subgroups of Chemaxon descriptors are elemental analysis, charge, geometry, partitioning, protonation, and isomers that are generated at the specified pH value 7.4. For deriving global nano-SARs, all the three types of descriptors were considered. And as for species specific models the selection of descriptors was narrowed down to the ALogPS and Chemaxon descriptors in order to allow for easier and better understanding of the underlying toxicity mechanisms with the assistance of the descriptors.

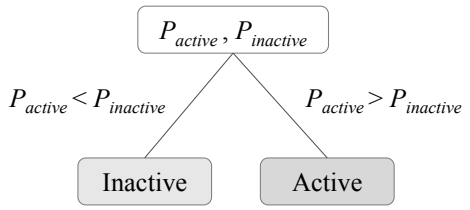


Figure 4.1. Decision test in a leaf node of a functional tree. P_{active} and $P_{inactive}$ are the categorical possibilities, f_{active} and $f_{inactive}$ are the regressions of input descriptors. Samples will be assigned to the group with the higher categorical possibility

4.2.3 Modeling algorithms

In order to build transparent rule-based nano-SARs that are easy to interpret and are capable of revealing information insight into the roles of structural descriptors, tree algorithms in Weka (version 3.6) were considered in the study. To avoid coincidence and also compare model performance, four extensively employed tree methods were used including functional tree, C4.5 decision tree, random tree, and simple CART (Hall et al., 2009).

In a functional tree model, both decision nodes and leaf nodes could contain tests based on either original input descriptors or the logistic regressions of descriptors (Gama, 2004). For binary classifications, prediction in the leaf nodes using logistic regressions of descriptors

could be explained as in Figure 4.1, where P_{active} and P_{inactive} are categorical possibilities needed to be compared; f_{active} and $f_{inactive}$ are the regressions of descriptors generated by the algorithm; inactive and active are the class labels to be returned for an observation.

The C4.5 decision tree is an extension of the earlier ID3 algorithm (Quinlan, 1986). It generates decision-based tree models in which each inner node contains a test only on the original input descriptors (Quinlan, 1993). For each test, a splitting cut-off value is provided and used for value comparison. The classification of ENM toxicity is accomplished by traversing a tree model from the root node to leaf nodes. Upon reaching the leaf nodes, labels (active or inactive) stored in the nodes will be returned as predictions.

The random tree algorithm constructs a tree randomly from a set of possible trees in which each tree has an equal chance of being sampled (Zhao and Zhang, 2008). A random tree is grown (without pruning) from data that has k randomly selected attributes at each node (Kukreja et al., 2012). The decision nodes contain queries only employing input descriptors and splitting thresholds, and leaf nodes comprise the category labels that an observation will be classified as. In the study, the k-value was set at 0 by default and the number of randomly chosen attributes was determined as $\log_2$(number of attributes) + 1. No depth restriction was set as the 'maxDepth' was 0 by default.

As a decision tree learner for classification, the simple CART (classification and regression tree) employs the minimal cost-complexity pruning of the CART algorithm when constructing predictive trees (Witten et al., 2011). It finds cost-complexity, a measure of average error reduced per leaf, and calculates the number of errors for each node when the subtrees are replaced by leaves (Rajput and Arora, 2013). The simple CART generates binary decision tree models for categorization issues. It handles the missing data by ignoring that record (Kalmegh, 2015).

4.2.4 Model performance evaluation

To estimate the predictive power of generated models, each dataset was randomly split into a training set (80%) and a test set (20%) before model construction. The learning process on the training set was executed in 10-fold cross validation to ensure the model stability. Predictive accuracy was characterized by four statistical parameters, defined as sensitivity (SE=TP/AP), specificity (SP=TN/AN), accuracy (Q=(TP+TN)/(AP+AN)), and correct classification rate (CCR=0.5(stability+ specificity)). Thereinto, TP represents the predicted number of true positives (or active class), TN stands for the predicted number of true negatives (or inactive class). AP and AN are numbers of actual positives and negatives observed, respectively. Reportedly, classification accuracy higher than 70% is considered as

high predictive performance (Kleandrova et al., 2014). And classification models with CCR of both training and test sets higher than 60.0% would be considered acceptable (Fourches et al., 2010). Model complexity was characterized by the size of the tree (number of nodes). Additionally, the significance of test sets was also verified by randomly permuting class labels of the test sets for global nano-SARs. The predictive results on these disjoint datasets should be approximately 50% (close to the no-information rate) for binary classifications with balanced datasets (Furlanello et al., 2003).

4.3 Results and discussion

4.3.1 Global nano-SARs across species

The influence of cut-off thresholds on model performances was primarily studied using the datasets I, II, and III. As can be seen in Figure S4.1, both high (0.1 mg/L) and low (100.0 mg/L) threshold values were evidenced to result in biased predictions. The thresholds selected for dataset I (LC50), II (EC50), and III (MIC) are respective 1.0, 10.0, and 10.0 mg/L to discretize numeric values for the case studies. After data discretization, dataset I was found to contain 175 ENMs of the active class and 225 of the inactive class; dataset II consisted of 246 ENMs labeled as active and 204 labeled as inactive; and dataset III has 87 ENMs from the active group and 79 from the inactive group. Using the OCHEM platform, 107, 95 and 122 computational descriptors were obtained for the datasets I, II, and III, respectively. Different nano-SARs were derived based on the descriptors which were linked to the nanotoxicity by the functional tree, C4.5 decision tree, random tree, and simple CART algorithms. An overview of the generated classification models is given in Table 4.1, in terms of modeling method, size of tree, sub-dataset, sensitivity, specificity, accuracy, and CCR. More details of the developed nano-SARs can be found in the Supplemental Information.

For case study I, the learning process was executed on the basis of 320 ENMs in the training set, while models were validated on the test set comprising 80 ENMs. A cut-off value of 1.0 mg/L was applied to enable the derivation of nano-SARs. By comparison, functional tree, C4.5 decision tree, and simple CART generated tree models with relatively low complexity (size of tree are respective 1, 5, and 11). As shown in Table 4.1, the random tree model was observed to be larger with a tree size of 55. These nano-SARs applied to the training set yielded accuracies of 70.9% (functional tree), 71.6% (C4.5 decision tree), 70.6% (random tree), and 69.1% (simple CART). Except for the simple CART model which correctly predicted 68.8% of the observations from the test set, accuracies of the LC50-

related nano-SARs on the test set were all found to exceed 70.0%. The CCR values calculated on sensitivity and specificity are higher than 60.0% for all the four models. Specifically, the C4.5 decision tree model merely contains two structural descriptors maximalprojectionsize and molecularpolarizability which belong to the Chemaxon descriptors. The descriptor maximalprojectionsize relates to the size of the molecule perpendicular to the minimal projection area surface (based on the van der Waals radius). And molecularpolarizability associates with the polarizability of the molecule. This means that the influence of both size and polarizability of the core element of ENMs was indicated. The simple CART model consists of five descriptors correlated with the geometrical size (minimalprojectionsize, maximalprojectionarea, minimalprojectionradius), molecular polarizability (averagemolecularpolarizability), and accessible surface areas of all atoms with negative partial charge (asa_ASA-). Owing the higher model complexity, however, the simple CART model was found to yield no higher predictive performance compared to the C4.5 decision tree. The functional tree has a relatively simpler tree structure with only one node but used more input descriptors in the logistic regressions.

With respect to the case study II, the 450 ENMs were randomly distributed to a training set of 360 ENMs and a test set of 90 ENMs. Numeric values of EC50 were discretized by a threshold of 10 mg/L. ENMs with EC50 values less than 10 mg/L were labeled as active, and the rest of ENMs were considered inactive. From the results show in Table 4.1, accuracies of all the models are between 60.0% and 65.0% for both training sets and test sets. This resulted from the low specificity of the nano-SARs while the models' sensitivities were considered reasonable. Thus the constructed EC50 models possess relatively low predictability for the inactive class. The unbalanced performances on both classes also resulted in the low CCRs between 60.0% and 65.0%.

Moreover, SAR-like models were also developed to predict the MICs of ENMs to various bacteria. In case study III, 133 ENMs were used to train the models and 33 ENMs were left out for the external validation. A threshold of 10.0 mg/L categorizes the ENMs into the active class (MIC $<$ 10.0 mg/L) or the inactive class (MIC $\geq$ 10.0 mg/L). The results depicted in Table 4.1 show that the C4.5 decision tree and the simple CART models exhibited the best predictability on the training set (both 75.9%), followed by the functional tree (75.2%) and the random tree models (70.7%). Predictive performance of the four nano-SARs on the test set gave the same results of 69.7% accuracy. CCRs of the training set are higher than 70.0% and those of the four test sets are all 69.7%. Except the most complex random tree model, the functional tree, C4.5 decision tree, and simple CART models have the same tree size of 3. Meanwhile, for both the C4.5 decision tree and the simple CART only the structural descriptor ALogPS_logS appeared in the built nano-SARs which is

associated with water solubility. The functional tree constructed the models using eight descriptors in its logistic regressions as can be seen in the Supplemental Information.

Table 4.1. Classification performances of the derived nano-SARs in case study I, II, and III. FT - functional tree; C4.5 - C4.5 decision tree; RT - random tree; n_{traing} - number of ENMs in the training set; n_{test} - number of ENMs in the test set. Details of the selection of the threshold values was described in the Supplemental Infromation

Method	Size of tree	Dataset	Sensitivity	Specificity	Accuracy	CCR
Case study I – LC50 (n_{traing} = 320, n_{test} = 80), threshold value 1.0 mg/L						
FT	1	Training set	0.750	0.678	0.709	0.714
		Test set	0.686	0.733	0.713	0.710
C4.5	5	Training set	0.671	0.750	0.716	0.711
		Test set	0.686	0.733	0.713	0.710
RT	55	Training set	0.679	0.728	0.706	0.704
		Test set	0.629	0.778	0.713	0.704
Simple CART	11	Training set	0.707	0.678	0.691	0.693
		Test set	0.686	0.689	0.688	0.688
Case study II – EC50 (n_{traing} = 360, n_{test} = 90), threshold value 10.0 mg/L						
FT	1	Training set	0.741	0.503	0.633	0.622
		Test set	0.796	0.415	0.622	0.606
C4.5	9	Training set	0.695	0.546	0.628	0.621
		Test set	0.816	0.415	0.633	0.616
RT	39	Training set	0.741	0.479	0.622	0.610
		Test set	0.816	0.439	0.644	0.628
Simple CART	17	Training set	0.650	0.564	0.611	0.607
		Test set	0.796	0.439	0.633	0.618
Case study III – MIC (n_{traing} = 133, n_{test} = 33), threshold value 10.0 mg/L						
FT	3	Training set	0.743	0.762	0.752	0.753
		Test set	0.706	0.688	0.697	0.697
C4.5	3	Training set	0.743	0.778	0.759	0.761
		Test set	0.706	0.688	0.697	0.697
RT	13	Training set	0.814	0.587	0.707	0.701
		Test set	0.706	0.688	0.697	0.697
Simple CART	3	Training set	0.743	0.778	0.759	0.761
		Test set	0.706	0.688	0.697	0.697

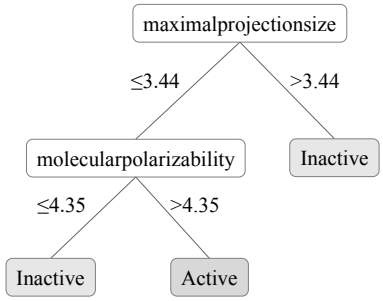


Figure 4.2. Developed C4.5 decision tree for the LC50 of metal-based ENMs. If LC50 < 1 mg/L the ENM is judged as active, and if LC50 ≥ 1mg/L the ENM is inactive.

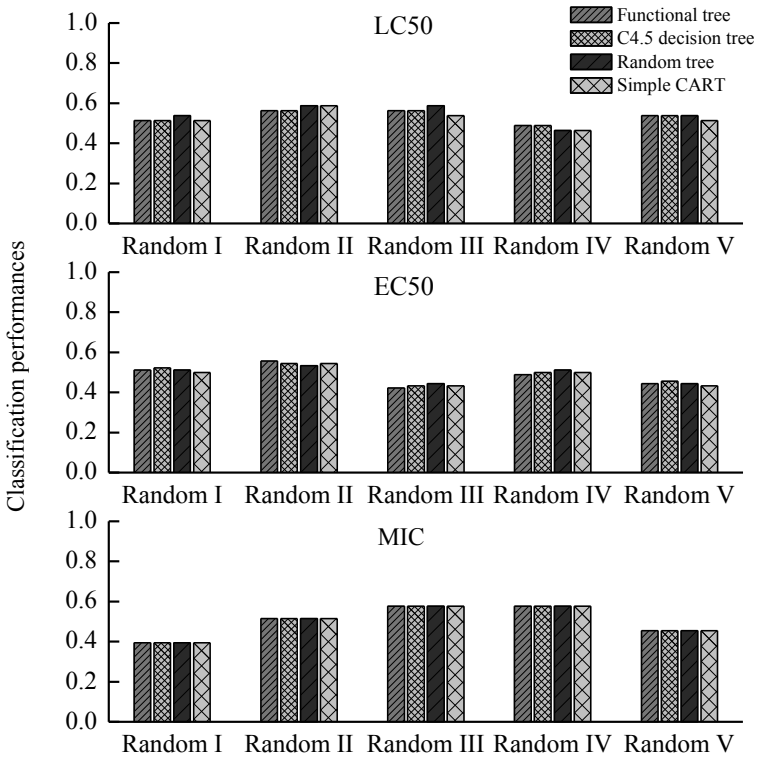
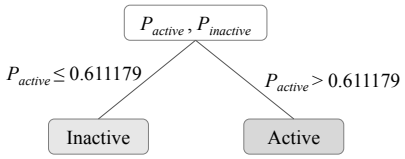
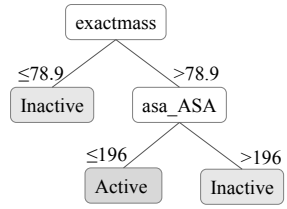


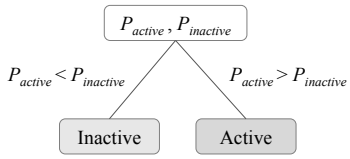
Figure 4.3. Model classification performances on randomized test sets. To verify the significance of the test sets of the three case studies, class labels in each test set were permuted for five times which yielded the randomized test sets Random I, II, III, IV, and V. For binary classifications, accuracy of the models on these disjoint test sets should be approximately 50% (the no-information rate).

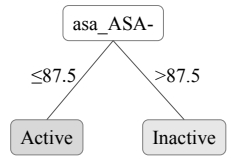




Danio rerio

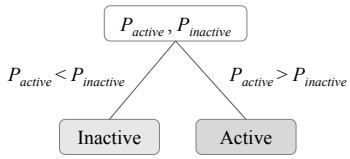
f_{active} = 0.33 + 0.19×averagemolecularpolarizability + 0.12×molecularpolarizability + 0.01×mass - 0.06×volume + 0.14×plattindex - 0.04×apKb1 - 0.58×ALogPS_logS = - $f_{inactive}$

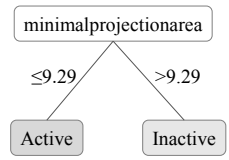




Daphnia magna

f_{active} = -0.36 + 0.1×molecularpolarizability - 0.16×tholepolarizability_a_xx - 0.51×tholepolarizability_a_zz + 0.01×exactmass + 0.01×volume - 0.38×logp + 0.01×asa_ASA+ - 0.01×asa_ASA_P = -$f_{inactive}$





Pseudokirchneriella subcapitata

f_{active} = 1.66 + 0.29×molecularpolarizability - 0.12×tholepolarizability_a_yy + 0.01×mass - 0.28×minimalprojectionarea - 0.01×volume - 0.01×dreidingenergy - 0.36×hyperwienerindex + 0.59×ALogPS_logS = -$f_{inactive}$

Figure 4.4. Developed functional tree (left) and C4.5 decision tree (right) models for *Danio rerio* (fish), *Daphnia magna* (crustacean), and *Pseudokirchneriella subcapitata* (algae). For the functional tree nano-SARs, P_{active} and P_{inactive} can be calculated as $P_{active}=\frac{\exp(f_{active})}{\exp(f_{active})+\exp(f_{inactive})}$, $P_{inactive}=\frac{\exp(f_{inactive})}{\exp(f_{active})+\exp(f_{inactive})}$.

The LC50 related functional tree, C4.5 decision tree, and random tree models showed reasonable predictability with accuracy (on training and test sets) higher than 70.0% and CCR higher than 60.0%, and with balanced performances on both categories. Based on a training set of 320 ENMs and test set of 80 ENMs, the C4.5 decision tree model is seen as relatively more concise as it only contains 5 nodes in the tree and uses two structural descriptors (maximalprojectionsize and molecularpolarizability), as shown in Figure 4.2.

Models presented in case study III were also considered acceptable based on the sensitivity, specificity, accuracy, CCR, and also tree complexity. As the developed nano-SARs exhibited similar predictive results on test sets, the significance of the test sets used in external validation was subsequently examined. We permuted the class labels in each test set for five times and validated the models with these randomized datasets afterwards. The results are depicted in Figure 4.3. As to case study I, the predictive accuracies on permuted test sets are between 46.3% and 58.8%. For case study II and III, it is 42.2% - 55.6 and 39.4% - 57.6%, respectively. Thus for all three cases, performances of the developed nano-SARs on the disjoint datasets are approximately 50% which is close to the no-information rate for binary classifications (Furlanello et al., 2003). It is therefore concluded that the original test sets are significant for model validation in the case studies I, II, and III.

Table 4.2. Performance of species-specific nano-SARs with the statistically significant predictability. FT - functional tree; C4.5 - C4.5 decision tree; n_{traing} - number of ENMs in the training set; n_{test} - number of ENMs in the test set

	Threshold (mg/L)	**Dataset**	**Sensitivity**	**Specificity**	**Accuracy**	**CCR**
Danio rerio, $n_{training}$ = 76, n_{test} = 18, LC50						
FT	100	Training set	0.943	0.913	0.934	0.928
		Test set	1.000	1.000	1.000	1.000
C4.5		Training set	0.906	0.913	0.908	0.910
		Test set	1.000	1.000	1.000	1.000
Daphnia magna, $n_{training}$ = 82, n_{test} = 20, LC50						
FT	1	Training set	0.843	0.968	0.890	0.906
		Test set	0.750	1.000	0.850	0.875
C4.5		Training set	0.843	0.968	0.890	0.906
		Test set	0.750	1.000	0.850	0.875
Pseudokirchneriella subcapitata, $n_{training}$ = 53, n_{test} = 13, EC50						
FT	1	Training set	0.944	0.914	0.925	0.929
		Test set	0.750	1.000	0.923	0.875
C4.5		Training set	0.944	0.914	0.925	0.929
		Test set	0.750	1.000	0.923	0.875
Staphylococcus aureus, $n_{training}$ = 32, n_{test} = 7, MIC						
C4.5	100	Training set	0.833	0.875	0.844	0.854
		Test set	0.800	1.000	0.857	0.900

4.3.2 Species-specific nano-SARs

Besides global models, species-specific nano-SARs were also built using the retrieved experimental data. This is in accordance with the recommendation of EU Directive 93/67/EEC ranking the hazards of ENMs to aquatic species. To begin with, from each dataset two species with the most data records were chosen for model development, which are *Danio rerio* (94 records) and *Daphnia magna* (102 records) from dataset I, *Daphnia magna* (105 records) and *Pseudokirchneriella subcapitata* (66 records) from dataset II, and *Escherichia coli* (41 records) and *Staphylococcus aureus* (39 records) from dataset III. For building models, two typical tree algorithms among the four selected methods, the functional tree and C4.5 decision tree algorithms were employed along with the ALogPS and Chemaxon descriptors. Cut-off thresholds investigated are respective 1, 10, and 100 mg/L. Performances of the derived nano-SARs are summarized in Table S4.1, Table S4.2, and Table S4.3 in the Supplemental Information. Models that exhibited significant predictive power are summarized and described in Table 4.2 and Figure 4.4. Nano-SARs were obtained for different hierarchies of species, i.e. *Danio rerio* (fish), *Daphnia magna* (crustacean), *Pseudokirchneriella subcapitata* (algae), and *Staphylococcus aureus* (bacteria). Details of these nano-SARs are presented in Table 4.3, including the number of ENMs in the training and test sets, size of the developed tree model, number of descriptors, and the names of descriptors involved.

The nano-SARs categorizing nanotoxicity to *Danio rerio* gave accuracies of 93.4% (functional tree) and 90.8% (C4.5 decision tree) on corresponding training sets (76 ENMs), and 100% accuracy on the two test sets (18 ENMs). Sensitivity and specificity of the two models are all above 90.0% on the training and test set (Table 4.2). This demonstrates the high predictability of the developed models. Model stability was ensured by executing 10-fold cross validation. Size of the corresponding functional tree model is 3 which means the nano-SAR only consists of one inner node and two decision nodes. As to *Daphnia magna*, the training set has 82 ENMs as samples for the learning process and the test set is comprised by 20 ENMs for validation. Accuracies of both the functional tree and the C4.5 decision tree models were shown to be 89.0% (training set) and 85.0% (test set) that are statistically significant. The CCRs of the model exceeded 85.0%. As shown in Table 4.3, the sizes of the functional tree and the C4.5 decision tree are respectively 1 and 3. With regards to *Pseudokirchneriella subcapitata*, functional tree and C4.5 decision tree models were built on the basis of 53 ENMs and validated by 13 ENMs. Predictive accuracies are as high as 92.5% on training set and 92.3% on test set with regard to both the functional tree and C4.5 decision tree with high CCR values. Moreover, built on a training set of 32 ENMs, the C4.5 decision tree model predicting the MIC to *Staphylococcus aureus* also exhibited significant predictability of 84.4% and 85.7% for the training and test set, respectively.

Table 4.3. Descriptor details of the species-specific nano-SARs. FT - functional tree; C4.5 - C4.5 decision tree

Nano-SAR	Method	ENMs number	Tree size	Descriptor number	List of descriptors
Danio rerio LC50 values	FT	94	3	7	averagemolecularpolarizability, molecularpolarizability, mass, volume, plattindex, apKb1, ALogPS_logS
	C4.5	94	5	2	exactmass, asa_ASA
Daphnia magna LC50 values	FT	94	1	8	molecularpolarizability, tholepolarizability_a_xx, tholepolarizability_a_zz, exactmass, volume, logp, asa_ASA+, asa_ASA_P
	C4.5	94	3	1	asa_ASA-
Pseudokirchneriella subcapitata EC50 values	FT	66	1	8	molecularpolarizability, tholepolarizability_a_yy, mass, minimalprojectionarea, volume, dreidingenergy, hyperwienerindex, ALogPS_logS
	C4.5	66	3	1	minimalprojectionarea
Staphylococcus aureus MIC values	C4.5	39	3	1	ALogPS_logS

Notably, even though mechanisms of the toxicity induced by metal-based ENMs to various hierarchies of species may vary, some descriptors in the models characterizing similar factors of ENMs were commonly observed and identified. As shown in Table 4.3 and Figure 4.4, descriptors representing molecular polarizability frequently appeared in the functional tree models. Those descriptors include the averagemolecularpolarizability, molecularpolarizability, molecularpolarizability, tholepolarizability_a_xx, tholepolarizability_a_zz, molecularpolarizability, and tholepolarizability_a_yy, which characterize different aspects of the electronic polarizability's contribution to nanotoxicity. Molecular polarizability measures the ability of the outer shell electrons in a molecule to move easily toward an external perturbation (Katritzky et al., 2007). Higher polarizability of the electrons in a molecule results in easier movement of electrons induced by an external electric field, which may trigger a series of biological reactions and lead to the toxicity of the materials (Singh and Gupta, 2014). For instance, detachment of an electron activated by solar radiation could stimulate the generation of hydroxyl radical $OH^{\bullet}$ as described in the study of Kar et al. (2014):

$$TiO_2 \overset{h\nu}{\Rightarrow} TiO_2^+ + \bar{e}$$

$$\bar{e} + O_2 \rightarrow O_2^{\bullet -}$$

$$O_2^{\bullet -} + 2H^+ + \bar{e} \rightarrow H_2O_2$$

$$O_2^{\bullet -} + H_2O_2 \rightarrow OH^{\bullet} + OH^- + O_2$$

$$H^+ + H_2O \Rightarrow OH^{\bullet} + H^+$$

Another discriminating factor is the accessible surface area of ENM cores that is quantified by asa_ASA (solvent accessible surface area), asa_ASA+ (solvent accessible surface area of all atoms with positive partial charge), asa_ASA_P (solvent accessible surface area of all polar atoms), and asa_ASA- (solvent accessible surface area of all atoms with negative partial charge) in the nano-SARs. The accessible surface area is defined as the accessible surface of molecules to a solvent (Zhang et al., 2008). For uncoated ENMs, the exposed surface area to the surroundings reflects the amount of atoms to be displayed on the surface and the potential of molecules to interact with the subcellular structures of species. As acknowledged, one of the outstanding properties of ENMs is the higher surface/volume ratio compared to that of their bulk counterparts which provides them increased surface reactivity and therefore possibly high toxicity (Li et al., 2008). As surface coatings are able to influence the toxicity of ENMs to species, surface area of ENM core still seems to play a role in nanotoxicity for the ENMs with modified surface. Moreover, descriptors quantifying the solubility were also observed such as apKb1 (dissociation constant) and ALogPS_logS (solubility in water) generated by OCHEM. Previous studies have shown that ENMs with less hardness and high solubility tend to exhibit stronger hazard effects (Gajewicz et al., 2015). This may be because the metal-ion leaching from ENM surface could act as one of the key factors inducing nanotoxicity (Hua et al., 2014; Xiao et al., 2015). Take Cu ENMs as an example, the release of Cu^{2+} from Cu-based nanoparticles could cause the generation of $OH^{\bullet}$ as follows (Stohs and Bagchi, 1995):

$$O_2^{\bullet -} + Cu^{2+} \rightarrow O_2 + Cu^+$$

$$Cu^+ + H_2O_2 \rightarrow Cu^{2+} + OH^- + OH^{\bullet}$$

The toxicity of ENMs may occur when the derived reactive oxygen species and the ions *per se* jointly or independently interact with the subcellular structures of species. Meanwhile, the geometrical descriptors minimalprojectionarea and minimalprojectionarea were also utilized in the model which indicate the spatial arrangement of the atoms forming a molecule. These

descriptors are associated with the molecular surface information obtained from atomic van der Waals areas and their overlap (Singh and Gupta, 2014). The descriptors relate to mass (mass, exactmass) and complexity (plattindex) were used in the nano-SARs as well. The Platt index is the sum of the degrees of all edges in the molecular graph, and is a considerably better measure of molecular complexity than merely the number of edges (Balaban et al., 1983; Saitta and Zucker, 2013).

4.3.3 Implications to the risk assessment of ENMs

On the basis of the computational descriptors offered by OCHEM and the assembled ecotoxicity data of metal-based nanomaterials, the developed LC50- and MIC-related global models and the species-specific nano-SARs showed reasonable predictive power. This demonstrates that it is indeed feasible to build nano-SARs using multi-sources datasets if the structures of ENMs are appropriately characterized. It also again confirms that the nano-(Q)SARs ought to be viewed as a potentially helpful tool in assisting the expedited review of ENM hazard categorization for the risk assessment of nanomaterials. With the experimental data retrieved from different scientific resources inconsistently characterizing the structures of ENMs, we managed to build nano-SARs classifying the nanoecotoxicity using descriptors solely representing the ENM cores. Such modeling tasks employing large datasets critically rely on the availability and quality of the datasets, and also on the comprehensive representation of ENM structures based on provided information. To accelerate the development of (Q)SAR-like models for nanomaterials much needs to be improved. Agreement on better data quality and availability are essential for nano-(Q)SARs with respect to both the toxicological and the componential aspects of the studied ENMs (Winkler, 2016). That is, the problem so far of the successful application of computational nanotoxicology is rather experimental, together with inadequate computational quantifications of ENM structures, than mathematical or statistical (Winkler, 2016). Unlike individual chemicals that are structurally unambiguous and possibly less complex, nanomaterials often exist as populations of materials varying in sizes, shapes, composites, and functional groups, etc. which can all significantly influence their biological interactions with environmental species (Chen et al., 2015). The structural uncertainty of the materials brings difficulty to experimentalists to offer complete and precise characterization of ENM structures, which subsequently hinders the calculation of representative descriptors for ENMs even when the compositions may have been properly provided (Fourches et al., 2011). The lack of data consistency especially in characterizing the structure of ENMs prevents the use of experimental data in developing nano-(Q)SARs, and may be one of the driving reasons why only a few datasets have been repeatedly used by the state-of-art of nano-(Q)SARs.

4.4 Conclusions

In this study, global nano-SARs across species and species-specific models classifying the ecotoxicity of metal-based ENMs were proposed. The models are intended to assist the nanomaterial hazard categorization and facilitate the ENM-related risk assessment and regulatory decision-making. To test the availability of existing nanotoxicity data in developing nano-(Q)SARs, datasets containing ecotoxicity information of ENMs from various publications or scientific resources were used including the LC50 (400 ENMs), EC50 (450 ENMs), and MIC (166 ENMs) related datasets. Due to the limited information characterizing the coating and functional groups of ENMs, descriptors were generated by the OCHEM to represent the core of the metal-based ENMs. Using the tree algorithms selected, easily interpretable and applicable classification nano-SARs were derived with significant predictability. The LC50 and MIC related global nano-SARs exhibited up to more than 70% accuracy of classification. The species-specific models were also developed to categorize the toxicity of metal-based ENMs to *Danio rerio*, *Daphnia magna*, *Pseudokirchneriella subcapitata*, and *Staphylococcus aureus*. Descriptor analysis indicated the role of molecular polarizability, accessible surface area, and metal-ion leaching in affecting the ecotoxicity of ENMs.

Acknowledgements

We thank the WEKA Machine Learning Project for the open-source software, and the OCHEM Team for the free access of the OCHEM platform. Guangchao Chen greatly thanks the funding support by the Chinese Scholarship Council (201306060076). This research is funded by the NATO project number SFPP·984401. Martina Vijver is funded by NWO-VIDI project number 864.13.010.

References

Balaban AT, Motoc I, Bonchev D, Mekenyan O, in Steric Effects in Drug Design, ed. Charton M, Motoc I, Springer, Berlin Verlag, 1983, Topological indices for structure-activity correlations, 21-55.

Chen G, Vijver MG, Peijnenburg WJ. Summary and analysis of the currently existing literature data on metal-based nanoparticles published for selected aquatic organisms: Applicability for toxicity prediction by (Q)SARs. Altern Lab Anim. 2015, 43:221-40.

Commission of the European Communities (CEC), Technical Guidance Document in Support of Commission Directive 93/67/EEC on Risk Assessment for New Notified Substances. Part II, Environmental Risk Assessment. Luxembourg, Luxembourg: Office for Official Publications of the European Communities, 1996.

Fourches D, Pu D, Tassa C, Weissleder R, Shaw SY, Mumper RJ, Tropsha A. Quantitative nanostructure-activity relationship modeling. ACS Nano. 2010, 4:5703-12.

Fourches D, Pu D, Tropsha A. Exploring quantitative nanostructure-activity relationships (QNAR) modeling as a tool for predicting biological effects of manufactured nanoparticles. Comb Chem High Throughput Screen. 2011, 14:217-25.

Furlanello C, Serafini M, Merler S, Jurman G. Entropy-based gene ranking without selection bias for the predictive classification of microarray data. BMC Bioinformatics, 2003, 4:54.

Gajewicz A, Schaeublin N, Rasulev B, Hussain S, Leszczynska D, Puzyn T, Leszczynski J. Towards understanding mechanisms governing cytotoxicity of metal oxides nanoparticles: Hints from nano-QSAR studies. Nanotoxicology. 2015, 9:313-25.

Gama J. Functional trees. Mach Learn. 2004, 55:219-250.

Godwin H, Nameth C, Avery D, Bergeson LL, Bernard D, Beryt E, Boyes W, Brown S, Clippinger AJ, Cohen Y, Doa M, Hendren CO, Holden P, Houck K, Kane AB, Klaessig F, Kodas T, Landsiedel R, Lynch I, Malloy T, Miller MB, Muller J, Oberdorster G, Petersen EJ, Pleus RC, Sayre P, Stone V, Sullivan KM, Tentschert J, Wallis P, Nel AE. Nanomaterial categorization for assessing risk potential to facilitate regulatory decision-making. ACS Nano. 2015, 9:3409-17.

Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. The WEKA data mining software: an update. SIGKDD Explor. 2009, 11:10-18.

Hua J, Vijver MG, Richardson MK, Ahmad F, Peijnenburg WJ. Particle-specific toxic effects of differently shaped zinc oxide nanoparticles to zebrafish embryos (Danio rerio). Environ Toxicol Chem. 2014, 33:2859-2868.

Juganson K, Ivask A, Blinova I, Mortimer M, Kahru A. NanoE-Tox: New and in-depth database concerning ecotoxicity of nanomaterials. Beilstein J Nanotechnol. 2015, 6:1788-804.

Kalmegh S. Analysis of WEKA data mining algorithm REPTree, Simple CART and RandomTree for classification of Indian news. Int J Innov Sci Eng Technol. 2015, 2:438-446.

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. 2014, 107:162-9.

Katritzky AR, Pacureanu L, Dobchev D, Karelson M. QSPR modeling of hyperpolarizabilities. J Mol Model. 2007, 13:951-63.

Kleandrova VV, Luan F, González-Díaz H, Ruso JM, Melo A, Speck-Planche A, Cordeiro MN. Computational ecotoxicology: Simultaneous prediction of ecotoxic effects of nanoparticles under different experimental conditions. Environ Int. 2014, 73:288-94.

Kukreja M, Johnston SA, Stafford P. Comparative study of classification algorithms for immunosignaturing data. BMC Bioinformatics. 2012, 13:139.

Li SQ, Zhu RR, Zhu H, Xue M, Sun XY, Yao SD, Wang SL. Nanotoxicity of TiO_2 nanoparticles to erythrocyte in vitro. Food Chem Toxicol. 2008, 46:3626-31.

Liu R, Zhang HY, Ji ZX, Rallo R, Xia T, Chang CH, Nel A, Cohen Y. Development of structure-activity relationship for metal oxide nanoparticles. Nanoscale. 2013, 5:5644-53.

Quinlan JR. Induction of Decision Trees. Mach Learn. 1986, 1:81-106.

Quinlan JR. C4.5: Programs for Machine Learning, Morgan Kaufmann, San Francisco, USA, 1993.

Rajput S, Arora A. Designing spam model-classification analysis using decision trees. Int J Comput Appl T. 2013, 75:6-12.

Regulation (EC) No 1272/2008 of the European Parliament and of the Council on classification, labelling and packaging of substances and mixtures. Official Journal of the European Union, L353, 2008, pp 1-1355.

Regulatory Cooperation Council – Nanotechnology Initiative (RCC-NI). Work element 2. Development of a classification scheme for nanomaterials regulated under the new substances programs of Canada and the United States, 2013, p 17.

Russell WMS, Burch RL. The Principles of Humane Experimental Technique, Methuen, London, 1959.

Saitta L, Zucker JD. Abstraction in Artificial Intelligence and Complex Systems, Springer, New York, 2013.

Singh KP, Gupta S. Nano-QSAR modeling for predicting biological activity of diverse nanomaterials. RSC Adv. 2014, 4:13215-13230.

Stohs SJ, Bagchi D. Oxidative mechanisms in the toxicity of metal ions. Free Radic Biol Med. 1995, 18:321-36.

Sushko I, Novotarskyi S, Körner R, Pandey AK, Rupp M, Teetz W, Brandmaier S, Abdelaziz A, Prokopenko VV, Tanchuk VY, Todeschini R, Varnek A, Marcou G, Ertl P, Potemkin V, Grishina M, Gasteiger J, Schwab C, Baskin II, Palyulin VA, Radchenko EV, Welsh WJ, Kholodovych V, Chekmarev D, Cherkasov A, Aires-de-Sousa J, Zhang

QY, Bender A, Nigsch F, Patiny L, Williams A, Tkachenko V, Tetko IV. Online chemical modeling environment (OCHEM): web platform for data storage, model development and publishing of chemical information. Comput Aided Mol Des. 2011, 25:533-54.

Tantra R, Oksel C, Puzyn T, Wang J, Robinson KN, Wang XZ, Ma CY, Wilkins T. Nano(Q)SAR: Challenges, pitfalls and perspectives. Nanotoxicology. 2015, 9:636-42.

Winkler D. Recent advances, and unresolved issues, in the application of computational modelling to the prediction of the biological effects of nanomaterials. Toxicol Appl Pharmacol. 2016, 299:96-100.

Witten I, Frank E, Hall M. Data Mining: Practical Machine Learning Tools and Techniques, 3rd Edition, Morgan Kaufmann, San Mateo, CA, 2011.

Xiao Y, Vijver MG, Chen G, Peijnenburg WJ. Toxicity and accumulation of Cu and ZnO nanoparticles in Daphnia magna. Environ Sci Technol. 2015, 49:4657-4664.

Zhang J, Gao X, Xu J, Li M. in Research in Computational Molecular Biology, ed. M. Vingron and L. Wong, Springer, Berlin Heidelberg, 2008, Rapid and Accurate Protein Side Chain Prediction with Local Backbone Information, 285-299.

Zhao Y, Zhang Y. Comparison of decision tree methods for finding active objects. Adv Space Res. 2008, 41:1955-1959.

Chapter 4 Supplemental Information

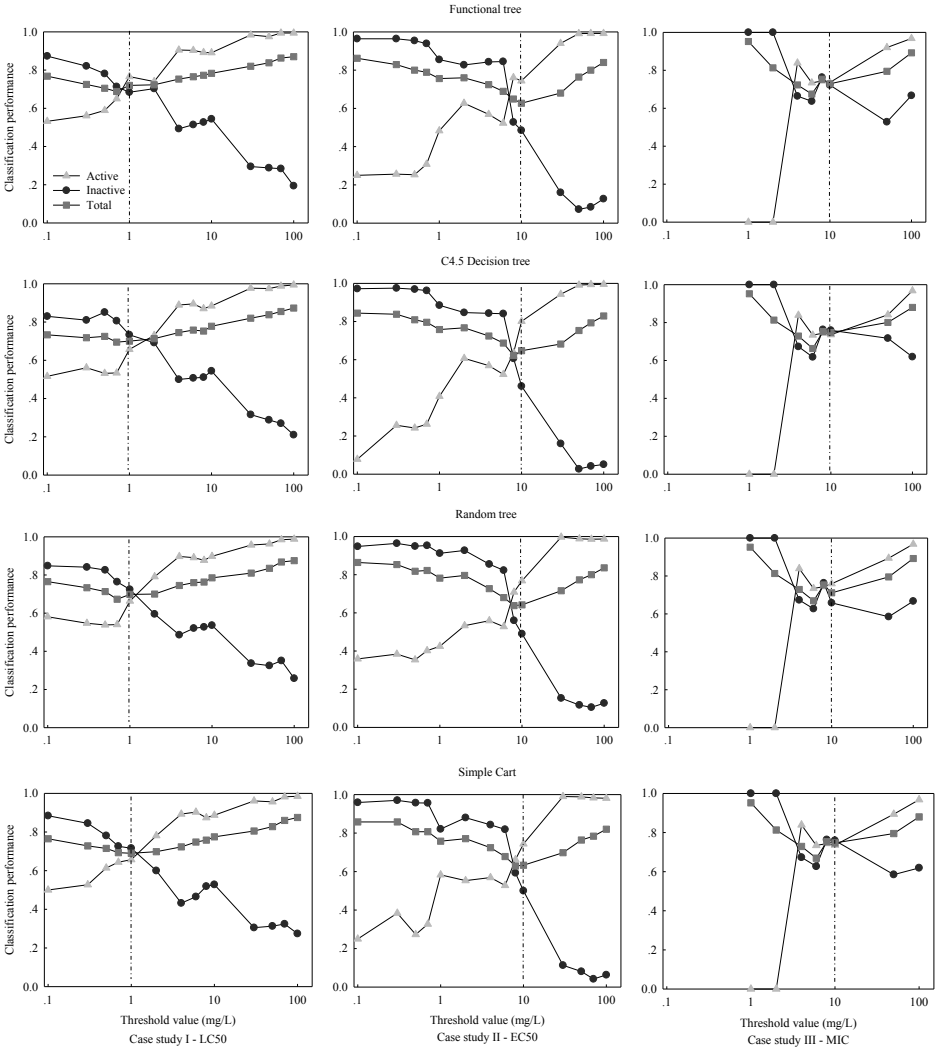


Figure S4.1. Effect of classification threshold values on model performances. Case study I: LC50 values of 400 ENMs, the optimal threshold value is 1 mg/L for discretizing the numerical values; Case study II: EC50 values of 450 ENMs, the optimal threshold value is 10 mg/L; Case study III: MIC values of 166 ENMs, the optimal threshold value is 10 mg/L.

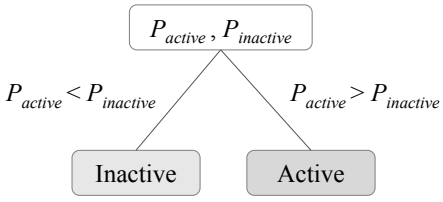
Influence of Discretization Thresholds on Model Performances

Before building models, the influences of discretization thresholds on model performances were taken into consideration. For global models, a series of thresholds were chosen to examine the tendency of model predictability with the shift of thresholds. Values of the thresholds were set to be 0.1, 0.3, 0.5, 0.7, 1.0, 2.0, 4.0, 6.0, 8.0, 10.0, 30, 50, 70, and 100.0 mg/L for the case studies of LC50 and EC50, and 1.0, 2.0, 4.0, 6.0, 8.0, 10.0, 50.0, and 100.0 mg/L for MIC due to a narrower variation of toxicity values. Within each dataset the records were ranked based on the values of toxicity endpoints. ENMs with toxicity values less than the threshold values were assigned to the 'active' class, and the rest of ENMs were labeled as 'inactive'. On the basis of different classification performances, the thresholds that lead to the most balanced predictive performances for both active and inactive groups were considered for the three case studies. Referring to the regulations and directives nowadays in force, choice of the thresholds for global models was restricted to the values of 0.1, 1.0, 10.0, and 100.0 mg/L, which are, for instance, used by the CLP-Regulation (EC) No 1272/2008 and the EU Directive 93/67/EEC to rank the hazard effects of chemicals. As results, selected cut-off values for case studies I, II, and III are respective 1, 10, and 10 mg/L, as described in Figure S1.

Developed Classification Models

Case study I LC50:

Functional tree



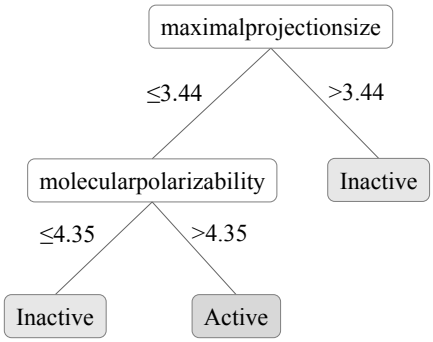
$$P_{active}=\frac{\exp(f_{active})}{\exp(f_{active})+\exp(f_{inactive})},\ P_{inactive}=\frac{\exp(f_{inactive})}{\exp(f_{active})+\exp(f_{inactive})}$$

f_{active} = 2.28 - 0.07×[tholepolarizability_a_zz] - 0.01×[volume] - 0.03×[polarsurfacearea] - 0.1×[SddTi] + 0.14×[SsAg] - 15.18×[SdAg] - 0.63×[Se1Al1Al1] - 0.34×[SsCo] - 5.56×[SdCa] - 0.26×[SsSn] + 0.37×[SsNi] - 0.21×[SsSe] + 1.48×[ALogPS_logP] = $-f_{inactive}$

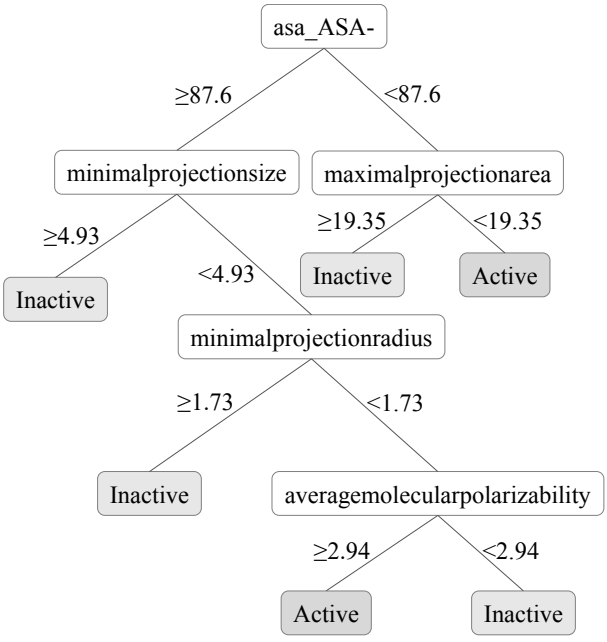
Random tree

```
asa_ASA_P< 78
|   minimalprojectionarea < 10.3
|   |   exactmass < 82.45
|   |   |   ALogPS_logS < 0.45
|   |   |   |   ALogPS_logS < 0.1 : Inactive (3/1)
|   |   |   |   ALogPS_logS >= 0.1 : Active (2/0)
|   |   |   ALogPS_logS >= 0.45 : Inactive (4/0)
|   |   exactmass >= 82.45
|   |   |   maximalprojectionradius < 3.01
|   |   |   |   wienerindex < 0.5
|   |   |   |   |   ALogPS_logP < -1.31
|   |   |   |   |   |   ALogPS_logS < 0.1
|   |   |   |   |   |   |   ALogPS_logS < 0.02 : Active (4/1)
|   |   |   |   |   |   |   ALogPS_logS >= 0.02 : Active (7/0)
|   |   |   |   |   |   ALogPS_logS >= 0.1
|   |   |   |   |   |   |   ALogPS_logS < 0.13 : Inactive (1/0)
|   |   |   |   |   |   |   ALogPS_logS >= 0.13
|   |   |   |   |   |   |   |   ALogPS_logS < 0.31 : Active (1/0)
|   |   |   |   |   |   |   |   ALogPS_logS >= 0.31
|   |   |   |   |   |   |   |   |   ALogPS_logS < 0.86
|   |   |   |   |   |   |   |   |   |   ALogPS_logS < 0.64 : Inactive (17/8)
|   |   |   |   |   |   |   |   |   |   ALogPS_logS >= 0.64 : Active (31/15)
|   |   |   |   |   |   |   |   |   ALogPS_logS >= 0.86 : Inactive (1/0)
|   |   |   |   |   ALogPS_logP >= -1.31 : Active (15/0)
|   |   |   |   wienerindex >= 0.5 : Active (18/8)
|   |   |   maximalprojectionradius >= 3.01
|   |   |   |   logd < -0.87 : Active (1/0)
|   |   |   |   logd>= -0.87 : Active (57/16)
|   minimalprojectionarea >= 10.3 : Inactive (8/0)
asa_ASA_P >= 78
|   asa_ASA- < 89.7
|   |   tholepolarizability_a_xx < 5.05
|   |   |   tholepolarizability_a_xx < 3.26 : Inactive (2/0)
|   |   |   tholepolarizability_a_xx >= 3.26
|   |   |   |   exactmass < 101.45
|   |   |   |   |   maximalprojectionradius < 2.45
|   |   |   |   |   |   minimalprojectionradius < 1.68 : Active (4/2)
|   |   |   |   |   |   minimalprojectionradius >= 1.68 : Inactive (40/12)
|   |   |   |   |   maximalprojectionradius >= 2.45
|   |   |   |   |   |   minimalprojectionradius < 1.67 : Inactive (4/0)
|   |   |   |   |   |   minimalprojectionradius >= 1.67 : Inactive (22/9)
|   |   |   |   exactmass >= 101.45 : Active (2/0)
|   |   tholepolarizability_a_xx >= 5.05 : Active (1/0)
|   asa_ASA- >= 89.7
|   |   molecularpolarizability < 6.67
|   |   |   chainbondcount < 1.5 : Inactive (3/0)
|   |   |   chainbondcount >= 1.5
|   |   |   |   minimalprojectionradius < 2.49
|   |   |   |   |   dreidingenergy < 61.85 : Inactive (2/0)
|   |   |   |   |   dreidingenergy >= 61.85
|   |   |   |   |   |   chainatomcount < 4 : Inactive (43/6)
|   |   |   |   |   |   chainatomcount >= 4 : Inactive (8/1)
|   |   |   |   minimalprojectionradius >= 2.49 : Inactive (8/2)
|   |   molecularpolarizability >= 6.67 : Inactive (11/0)
```

C4.5 decision tree

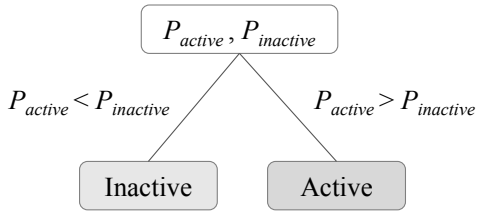


Simple CART



Case study II EC50:

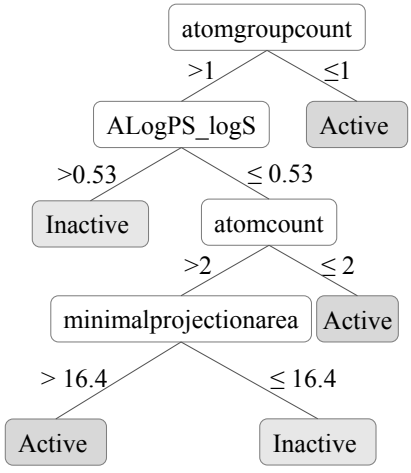
Functional tree



$$P_{active}=\frac{\exp(f_{active})}{\exp(f_{active})+\exp(f_{inactive})}, P_{inactive}=\frac{\exp(f_{inactive})}{\exp(f_{active})+\exp(f_{inactive})}$$

f_{active} = 0.46 - 0.22×[rotatablebondcount] - 0.11×[SsAg] - 0.35×[Se2Ni1O1] + 0.49×[Se1Au1Au1] - 0.44×[SdsDy] + 0.61×[Se1Er2O2ds] - 0.6×[SsFe] + 0.23×[SsAl] + 0.18×[SdsSb] = - $f_{inactive}$

C4.5 decision tree



Random tree

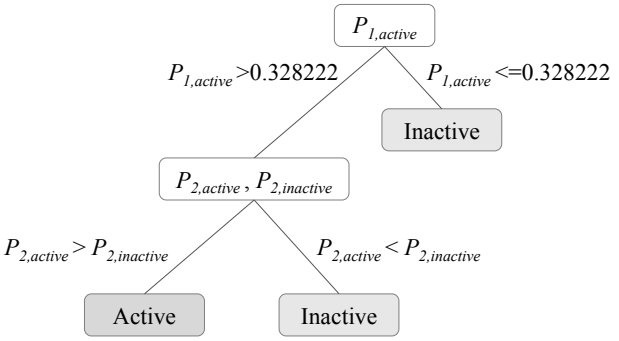
```
balabanindex < 1.32
|   molecularpolarizability < 4.33
|   |   maximalprojectionarea < 13.35
|   |   |   logp < -0.65
|   |   |   |   minimalprojectionarea < 7.07 : Inactive (1/0)
|   |   |   |   minimalprojectionarea >= 7.07 : Active (66/31)
|   |   |   logp >= -0.65
|   |   |   |   minimalprojectionsize < 4.73 : Active (59/29)
|   |   |   |   minimalprojectionsize >= 4.73 : Active (2/0)
|   |   maximalprojectionarea >= 13.35 : Inactive (5/0)
|   molecularpolarizability >= 4.33
|   |   minimalprojectionsize < 6.45
|   |   |   maximalprojectionsize < 0.9
|   |   |   |   ALogPS_logS < -0.03 : Active (16/6)
|   |   |   |   ALogPS_logS >= -0.03
|   |   |   |   |   ALogPS_logS < 0.7
|   |   |   |   |   |   ALogPS_logP < -1.11
|   |   |   |   |   |   |   ALogPS_logS < 0.31 : Active (3/0)
|   |   |   |   |   |   |   ALogPS_logS >= 0.31
|   |   |   |   |   |   |   |   ALogPS_logS < 0.54 : Active (34/4)
|   |   |   |   |   |   |   |   ALogPS_logS >= 0.54 : Active (1/0)
|   |   |   |   |   |   ALogPS_logP >= -1.11 : Active (8/2)
|   |   |   |   |   ALogPS_logS >= 0.7
|   |   |   |   |   |   ALogPS_logS < 0.95 : Active (24/8)
|   |   |   |   |   |   ALogPS_logS >= 0.95 : Active (3/0)
|   |   |   maximalprojectionsize >= 0.9
|   |   |   |   maximalprojectionradius < 3.19 : Active (12/0)
|   |   |   |   maximalprojectionradius >= 3.19 : Active (18/4)
|   |   minimalprojectionsize >= 6.45 : Inactive (1/0)
balabanindex >= 1.32
|   logd < -0.43
|   |   atomcount < 4
|   |   |   molecularsurfacearea < 88.1
|   |   |   |   minimalprojectionradius < 2.25 : Inactive (68/22)
|   |   |   |   minimalprojectionradius >= 2.25 : Inactive (3/0)
|   |   |   molecularsurfacearea >= 88.1 : Inactive (26/11)
|   |   atomcount >= 4 : Active (2/0)
|   logd >= -0.43 : Inactive (8/0)
```

Simple CART

```
asa_ASA_H < 61.0
|  minimalprojectionsize < 4.93
|  |  minimalprojectionarea < 7.029999999999999: Active(2.0/0.0)
|  |  minimalprojectionarea >= 7.029999999999999
|  |  |  averagemolecularpolarizability < 2.94: Active(35.0/31.0)
|  |  |  averagemolecularpolarizability >= 2.94: Active(30.0/29.0)
|  minimalprojectionsize >= 4.93
|  |  averagemolecularpolarizability < 7.52
|  |  |  tholepolarizability_a_xx < 3.4349999999999996
|  |  |  |  minimalprojectionarea < 12.55
|  |  |  |  |  tholepolarizability_a_yy < 4.4: Inactive(15.0/11.0)
|  |  |  |  |  tholepolarizability_a_yy >= 4.4: Inactive(46.0/22.0)
|  |  |  |  minimalprojectionarea >= 12.55: Inactive(3.0/0.0)
|  |  |  tholepolarizability_a_xx >= 3.4349999999999996: Inactive(14.0/0.0)
|  |  averagemolecularpolarizability >= 7.52: Active(2.0/0.0)
asa_ASA_H >= 61.0: Active(95.0/25.0)
```

Case study III MIC:

Functional tree



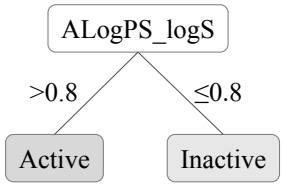
$$P_{1,active}=\frac{\exp(f_{1,active})}{\exp(f_{1,active})+\exp(f_{1,inactive})}$$

$$P_{2,active}=\frac{\exp(f_{2,active})}{\exp(f_{2,active})+\exp(f_{2,inactive})}, P_{2,inactive}=\frac{\exp(f_{2,inactive})}{\exp(f_{2,active})+\exp(f_{2,inactive})}$$

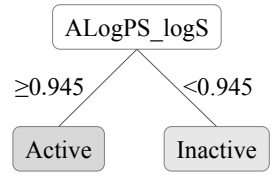
$f_{1,active}$ = 5.1 + 0.14×[minimalprojectionarea] + 0.01×[asa_ASA_H] - 1.07×[balabanindex] - 0.14×[hararyindex] - 0.01×[asa_ASA+] - 0.61×[SsCu] + 5.91×[ALogPS_logP] + 0.91×[ALogPS_logS] = -$f_{1,inactive}$

$f_{2,active}$ = 59.21 - 0.29×[averagemolecularpolarizability] + 0.14×[minimalprojectionarea] + 0.01×[asa_ASA_H] -1.07×[balabanindex] -0.14×[hararyindex] - 0.01×[asa_ASA+] - 0.61×[SsCu] + 46.68×[ALogPS_logP] + 0.39×[ALogPS_logS] = -$f_{2,active}$

C4.5 decision tree



Simple CART



Random tree

```
tholepolarizability_a_yy < 2.05
|   ALogPS_logP < -1.31
|   |   ALogPS_logS < 0.11 : Inactive (5/0)
|   |   ALogPS_logS >= 0.11
|   |   |   ALogPS_logS < 0.31 : Inactive (27/13)
|   |   |   ALogPS_logS >= 0.31 : Inactive (13/4)
|   ALogPS_logP >= -1.31 : Active (66/14)
tholepolarizability_a_yy >= 2.05
|   maximalprojectionarea < 21 : Inactive (18/0)
|   maximalprojectionarea >= 21
|   |   chainatomcount < 3.5 : Inactive (3/1)
|   |   chainatomcount >= 3.5 : Inactive (1/0)
```

Table S4.1. Performances of the LC50 related nano-SARs for *Danio rerio* and *Daphnia magna*. The best performance of the models were bolded in the table

	Threshold (mg/L)	**Data set**	**Sensitivity**	**Specificity**	**Accuracy**	**CCR**
Danio rerio, $n_{training}$ = 76, n_{test} = 18						
Functional tree	1	Training set	0.389	0.828	0.724	0.609
		Test set	0.750	0.714	0.722	0.732
	10	Training set	0.868	0.632	0.750	0.750
		Test set	0.667	0.556	0.611	0.612
	100	Training set	**0.943**	**0.913**	**0.934**	0.928
		Test set	**1**	**1**	**1**	1
C4.5 decision tree	1	Training set	0.056	0.948	0.737	0.502
		Test set	0	1	0.778	0.500
	10	Training set	0.947	0.632	0.789	0.790
		Test set	1	0.556	0.778	0.778
	100	Training set	**0.906**	**0.913**	**0.908**	0.910
		Test set	**1**	**1**	**1**	1
Daphnia magna, $n_{training}$ = 82, n_{test} = 20						
Functional tree	1	Training set	**0.843**	**0.968**	**0.890**	0.906
		Test set	**0.750**	**1**	**0.850**	0.875
	10	Training set	0.971	0.250	0.866	0.611
		Test set	0.941	0.333	0.850	0.637
	100	Training set	1	0	0.927	0.500
		Test set	0.947	0	0.900	0.474
C4.5 decision tree	1	Training set	**0.843**	**0.968**	**0.890**	0.906
		Test set	**0.750**	**1**	**0.850**	0.875
	10	Training set	0.957	0.250	0.854	0.604
		Test set	0.941	0.333	0.850	0.637
	100	Training set	1	0	0.927	0.500
		Test set	1	0	0.950	0.500

Table S4.2. Performances of the EC50 related nano-SARs for *Daphnia magna* and *Pseudokirchneriella subcapitata*. The best performance of the models were bolded in the table

	Threshold (mg/L)	**Data set**	**Sensitivity**	**Specificity**	**Accuracy**	**CCR**
Daphnia magna, $n_{training}$ = 84, n_{test} = 21						
Functional tree	1	Training set	0.552	0.909	0.738	0.731
		Test set	0.500	1	0.762	0.750
	10	Training set	0.926	0.313	0.810	0.620
		Test set	1	0.500	0.905	0.750
	100	Training set	1	0	0.929	0.500
		Test set	1	0	0.905	0.500
C4.5 decision tree	1	Training set	0.550	0.909	0.738	0.730
		Test set	0.500	1	0.762	0.750
	10	Training set	0.912	0.375	0.810	0.644
		Test set	0.824	0.750	0.810	0.787
	100	Training set	1	0	0.929	0.500
		Test set	1	0	0.905	0.500
Pseudokirchneriella subcapitata, $n_{training}$ = 53, n_{test} = 13						
Functional tree	1	Training set	**0.944**	**0.914**	**0.925**	0.929
		Test set	**0.750**	**1**	**0.923**	0.875
	10	Training set	0.813	0.667	0.755	0.740
		Test set	0.750	0.800	0.769	0.775
	100	Training set	1	0	0.906	0.500
		Test set	1	0	0.846	0.500
C4.5 decision tree	1	Training set	**0.944**	**0.914**	**0.925**	0.929
		Test set	**0.750**	**1**	**0.923**	0.875
	10	Training set	0.781	0.667	0.736	0.724
		Test set	0.750	0.800	0.769	0.775
	100	Training set	1	0	0.906	0.500
		Test set	1	0	0.846	0.500

Table S4.3. Performances of the MIC related nano-SARs for *Escherichia coli* and *Staphylococcus aureus*. The best performance of the models were bolded in the table

	Threshold (mg/L)	**Data set**	**Sensitivity**	**Specificity**	**Accuracy**	**CCR**
Escherichia coli, $n_{training}$ = 33, n_{test} = 8						
Functional tree	10	Training set	0	1	0.636	0.500
		Test set	0	1	0.625	0.500
	100	Training set	1	0	0.515	0.500
		Test set	0.882	0.563	0.727	0.723
C4.5 decision tree	10	Training set	0.250	0.905	0.667	0.578
		Test set	0	1	0.625	0.500
	100	Training set	0	1	0.5	0.500
		Test set	0.750	1	0.875	0.875
Staphylococcus aureus, $n_{training}$ = 32, n_{test} = 7						
Functional tree	10	Training set	1	0	0.563	0.500
		Test set	0.750	0.667	0.714	0.709
	100	Training set	1	0	0.750	0.500
		Test set	0.800	1	0.857	0.900
C4.5 decision tree	10	Training set	0.667	0.357	0.531	0.512
		Test set	0.750	0.667	0.714	0.709
	100	Training set	**0.833**	**0.875**	**0.844**	0.854
		Test set	**0.800**	**1**	**0.857**	0.900

CHAPTER 5

DEVELOPING SPECIES SENSITIVITY DISTRIBUTIONS FOR METALLIC NANOMATERIALS CONSIDERING THE CHARACTERISTICS OF NANOMATERIALS, EXPERIMENTAL CONDITIONS, AND DIFFERENT TYPES OF ENDPOINTS

Chen G, Peijnenburg WJGM, Xiao Y, Vijver MG

Published in *Food and Chemical Toxicology*. 2017, doi: 10.1016/j.fct.2017.04.003

Abstract

A species sensitivity distribution (SSD) for engineered nanomaterials (ENMs) ranks the tested species according to their sensitivity to a certain ENM. An SSD may be used to estimate the maximum acceptable concentrations of ENMs for the purpose of environmental risk assessment. To construct SSDs for metal-based ENMs, more than 1800 laboratory derived toxicity records of metallic ENMs from >300 publications or open access scientific reports were retrieved. SSDs were developed for the metallic ENMs grouped by surface coating, size, shape, exposure duration, light exposure, and different toxicity endpoints. It was found that PVP- and sodium citrate- coatings enhance the toxicity of Ag ENMs as concluded from the relevant SSDs. For the Ag ENMs with different size ranges, differences in behavior and/or effect were only observed at high exposure concentrations. The SSDs of Ag ENMs separated by both shape and exposure duration were all nearly identical. Crustaceans were found to be the most vulnerable group to metallic ENMs. In spite of the uncertainties of the results caused by limited data quality and availability, the present study provided novel information about building SSDs for distinguished ENMs and contributes to the further development of SSDs for metal-based ENMs.

Key words: ecotoxicity; engineered nanomaterial; modeling; risk assessment; species sensitivity distributions

5.1 Introduction

Over the last decade, products that incorporate nano-structured materials have been rapidly introduced to the market. In 2014, the value of the global market regarding nanotechnology products was estimated to be $26 billion, and is expected to reach about $65 billion by 2019 (Winkler, 2016). While the benefits of nanotechnology are beyond debate, the concern is increasing about the safe use and subsequent environmental impacts of engineered nanomaterials (ENMs). Evaluating the environmental risks of ENMs is essential to manage relevant risks and ensure the safety of these manufactured materials (Piperigkou et al., 2016; Toropova and Toropov, 2013). One of the well-established approaches assisting risk assessment of ENMs is the development of species sensitivity distributions (SSDs) (Gottschalk and Nowack, 2013). SSDs rank the species based on their sensitivity to a certain ENM, and reflect the potentially affected fraction of species under an exposure concentration of interest (Garner et al., 2015). From the SSD, among others the 5th percentile of the fitted distribution (HC5) can be derived. The HC5 is commonly used as the basis for environmental risk assessment of chemicals and is assumed to be the concentration that is sufficiently protecting ecosystems following addition of an extra safety factor that ranges in between 1 and 5 (European Chemicals Agency, 2008). Risk quantification is usually performed by dividing the predicted environmental concentration by either the predicted no observed effect concentration in case of specific species or by the HC5 in case of generic risk assessment. When the risk quotient is greater than or equals 1, a potential risk of the nanomaterials exists and further assessment is required, including the option of additional toxicity testing; when the risk quotient is less than 1, environmental risks are not expected.

Previously, a few examples of SSDs have been presented for different ENMs based on a limited set of laboratory derived toxicity data. To quantify the environmental risks of nano-Ag, nano-TiO_2, nano-ZnO, carbon nanotubes, and fullerenes in four environmental compartments (surface water, sewage treatment plant effluents, soils, and sludge-treated soils), SSDs were generated for the five ENMs (Gottschalk et al., 2013). The SSDs reflecting the no observed effect concentrations were then compared with the distributions of predicted environmental concentrations in the four environmental compartments. The results indicated marginal risks of Ag and TiO_2 ENMs to surface water species and a low level of risk caused by Ag, TiO_2, and ZnO ENMs in sewage treatment plant effluents. SSDs for the same five metallic ENMs were also generated by Coll et al. (2016) for different taxa. The risk quotients that are closest to 1 for both ZnO and TiO_2 ENMs among others indicated the highest priority of these materials to be studied in more depth. In another study, SSDs for seven types of metallic ENMs were built including Ag, Al_2O_3, CeO_2, Cu, CuO, TiO_2, and ZnO ENMs (Garner et al, 2015). The HC5 values with 95% confidence

interval (CI) of each ENM were calculated and compared with those of the corresponding ionic and bulk counterparts. The SSDs of PVP-coated and uncoated Ag ENMs were separately modeled, allowing to conclude about the influence of surface coatings on SSDs. As first attempts of developing SSDs for ENMs, those developed SSDs have provided significant information of the potential environmental impacts of ENMs, and contributed to the derivation of HC5 values as policy measures of the ENMs of concern. The further interest of the development of SSDs for ENMs would be, ideally, to cover more types of ENMs to comprehensively evaluate the risks of all the widely applied ENMs; and to include the large variety of environmental species in order to build robust and reliable SSDs. Meanwhile better estimates could be obtained when specific attention is paid in SSD development to specific ENM properties such as surface coating, size, and shape, and also to the dynamic behaviors of ENMs in the exposure media (Garner et al. 2015; Gottschalk et al., 2013). The consideration of ENM characteristics in developing SSDs may also provide hint messages for the safe-by-design of ENMs, if the SSDs of ENMs separated by certain characteristics were found to shift significantly compared with that separated by other properties. The implementation of the research needs mentioned here, is however strongly limited by the quality of published raw data from the ecotoxicity assays and to a lower extent by the limited availability of suited exposure and effect data.

In response to the above-mentioned challenges, the present study aims to investigate the availability of currently published ecotoxicity data of ENMs for their suitability in developing SSDs for metal-based ENMs; and secondly to build SSDs for ENMs considering the structural characteristics (e.g. surface coating, size, shape), experimental conditions, and also different types of toxicity endpoints. All together more than 1800 ecotoxicity records of metallic ENMs from >300 publications or open access scientific reports were retrieved from the databases of Chen et al. (2015), Juganson et al. (2015), and the online chemical modeling environment (OCHEM) (Sushko et al., 2011). The toxicity endpoints in the collected dataset include the lethal concentration (LC), the effect concentration at a specific effect level (ECx), the lowest observed effect concentration (LOEC), and the no observed effect concentration (NOEC). The studied species originated from seven widely investigated organism groups namely algae, bacteria, crustacean, fish, nematodes, protozoa, and yeast. Based on the analysis, the development of SSDs focuses on Ag, CeO_2, CuO, TiO_2, and ZnO ENMs due to relatively sufficient information availability. Different SSDs were generated for the Ag ENMs grouped by surface coating, size, shape, and exposure duration. The SSD for UV exposed TiO_2 ENMs was also derived. To determine whether and to what extent the shape of the SSD curve might alter and the HC5s may vary based on different toxicity endpoints, these topics were also considered in the development of SSDs in the present study. To discuss the vulnerability of different

organism groups and species to the metallic ENMs, the most sensitive species in each developed SSD was analyzed as well.

5.2 Methods

5.2.1 Datasets

Experimental data of ENM ecotoxicity were assembled from three databases. The first database is that developed by Chen et al. (2015) consisting of 886 records of toxicity endpoints of various metal-based ENMs. The second database is the NanoE-tox database listing in total 1518 EC50 (the concentration at which 50 % of the test species is affected), LC50 (median lethal concentration), and NOEC values regarding eight ENMs including carbon nanotubes and fullerenes, Ag, CeO_2, CuO, TiO_2, ZnO, and FeO_x nanomaterials (Juganson et al., 2015). The third data source is the OCHEM platform which explicitly provided 244 LC50 values and 170 EC50 values of different metallic ENMs (Sushko et al., 2011). After removing duplicate information, the newly developed dataset counts all together more than 1800 values of metallic ENMs from >300 publications or open access scientific reports. This information was afterwards filtered by the following conditions: a) toxicity of metal-based ENMs solely; b) tested organisms are algae, bacteria, crustacean, fish, nematodes, protozoa, and yeast only; c) toxicity endpoints are LC, EC, LOEC, and NOEC. In the dataset, units of all toxicity values were unified into mg/L, and the endpoints larger than 10000 mg/L were excluded as these are considered to be irrelevant from a toxicological point of view.

As for certain ENMs, the toxicity data was separated by the characteristics of the ENMs (i.e. surface coating, size, shape), experimental condition (duration of exposure, light exposure), and type of different endpoints (LC, EC, LOEC, NOEC), respectively. The number of species in each sub-dataset is required to be at least six in order to construct a reliable SSD (Cedergreen et al., 2004). SSDs for the uncoated and differently coated ENMs were modeled. With regard to grouping ENMs by size, it was suggested by Garner et al. (2015) to divide the data in size ranges in between 1-10, 10-50, and 50-100 nm. Here, we adapted the division of sizes as 1-20, 20-50, and 50-100 nm, as it was stated that nanoparticles with size <20 nm may have significantly increased surface reactivity and behave differently than larger particles (Auffan et al., 2009; 2010), whereas nanomaterials of 20-50 nm appear to be taken up more rapidly than particles of other sizes (Iversen, et al., 2011; Jin et al., 2009). When generating SSDs based on data separated by the size and shape, ENMs with reported surface coatings were excluded. The exposure duration was determined as ≤1 d, 1-2 d,

and >2 d, to investigate if over time the shape of SSD-curve might shift as result of both the dynamic changes of ENMs in the media and the increased length of the life cycle of an organism. The experimental condition of light exposure was also considered in the study as nanomaterials like TiO_2 ENMs were reportedly able to catalyze reactions under UV radiation and cause phototoxicity (Yin et al., 2012; Sanders et al., 2012).

5.2.2 Modeling algorithm

Data was grouped regarding LC50 value and ranked from lowest to highest by the following equation (US EPA, 1998):

$$\text{Proportion} = \frac{Rank - 0.5}{Number\ of\ species}$$

For the toxicity data relating sub-lethal effects of ENMs (i.e. EC50, LOEC, NOEC), the median toxicity values based on a certain biological effect to a species were initially calculated per reported effect. The obtained medians of different effects to that species were afterwards compared and the lowest median value was used in ranking the species sensitivities. The ranked median values of different species were then plotted against the cumulative probability which reflects the proportion of species affected at a certain concentration.

In the study, lognormal distributions of species sensitivity were fitted using the 'fitdistr' function of the MASS package in the R statistical software (version 3.3.1). This function generates a maximum-likelihood fitting of univariate distributions, allowing parameters to be held fixed if desired (Venables and Ripley, 2002). The 95% CI of the fitted regressions was also estimated by employing the strategy of parametric bootstrap. The HC5 values of the SSDs were extracted by the 'quantile' function in the R software (Hyndman and Fan, 1996).

5.3 Results

We firstly analyzed the data availability for the preparation of building SSDs for the metal-based ENMs (Table 5.1). Before constructing separate SSDs, the SSDs for Ag, CeO_2, CuO, TiO_2, and ZnO ENMs were generated with all available data for the corresponding ENMs (see Figure S5.1 as provided in the Supplemental Information). LOEC and NOEC data for Ag and CuO ENMs are available for only five species. These data were nevertheless

included in the analysis to allow for a more comprehensive comparison. Separate SSDs were afterwards obtained for Ag ENMs grouped by surface coating, size, shape, and exposure duration (Figure 5.1); for CuO and ZnO ENMs grouped by size (Figure S5.2); and for TiO_2 ENMs grouped by size and light exposure (Figure S5.2). SSDs based on different toxicity endpoints were compared (Figure 5.3). The significance of difference between relevant HC5s was discussed (Figure 5.2, Figure 5.4, and Figure S5.3). All the calculated HC5 values with corresponding CI were listed in a Microsoft Excel spreadsheet (see Supplemental Information). The lists of species that were used to build SSDs were also presented in the Supplemental Information. Examples of building SSDs in the present study using LC50, EC50, LOEC, and NOEC datasets were presented in the Supplemental Information.

Table 5.1. Number of species tested for Ag, CeO_2, Cu, CuO, Ni, TiO_2, ZnO, and other ENMs. The species are from seven groups of organisms, namely algae, bacteria, crustacean, fish, nematodes, protozoa, and yeast. ENMs with species number less than four (for every type of endpoint) are in the group 'Others'

ENMs	**LC50**	**EC50**	**LOEC**	**NOEC**
Ag	17	20	5	5
CeO_2	2	6	2	8
Cu	4	1	0	0
CuO	9	10	5	5
Ni	4	4	0	0
TiO_2	10	16	2	17
ZnO	8	13	6	11
Others	10	14	4	10

5.3.1 Data availability for generating SSDs

The information in the newly collected dataset includes but is not limited to: characteristics of ENMs (core, size, surface coating, shape, surface area etc.), experimental conditions (exposure duration, light exposure etc.), tested species, detected biological effects, type of toxicity endpoints, and values of nanotoxicity. The studied ENMs cover a wide range of types of ENMs such as Ag, CeO_2, CuO, FeO_x, NiO, SiO_2, TiO_2, ZnO ENMs etc. The toxicity endpoints that are potentially useful for building SSDs are LC50, EC50, LOEC, and NOEC, as data availability of other endpoints is very limited. In order to develop SSDs, the

number of species was analyzed for which data with regard to each type of ENMs and with respect to each type of the endpoint was available. The results of this analysis are shown in Table 5.1. ENMs for which data for each endpoint were available for no more than three species were included in the group 'Others'.

The analysis showed that Ag, CeO_2, Cu, CuO, Ni, TiO_2, and ZnO ENMs have received the most research attention among all the metallic ENMs. Ag ENMs have been shown to be generally studied for their lethal toxicity to different taxa (17 species), as well as its sub-lethal biological effects (20 species for which EC50 values were reported). CuO, TiO_2, and ZnO ENMs were also widely tested on various species, which provided toxicity data for respectively 9, 10, 8 species on LC50, and 10, 16, 13 species on EC50. For CeO_2 ENMs, 6 and 8 data points are available on EC50 and NOEC respectively. For Cu and Ni ENMs, the retrieved data for constructing SSDs is very limited based on both LC50 and EC50. Based on this analysis, we subsequently developed SSDs for the ungrouped Ag, CeO_2, CuO, TiO_2, and ZnO ENMs (Figure S5.1) and the ENMs differentiated by surface coating, size, shape, exposure duration, light exposure, and type of endpoint.

5.3.2 Separate SSDs by ENM characteristics and experimental conditions

Within the first constructed SSDs, uncoated, polyvinylpyrrolidone (PVP)- and sodium citrate- coated Ag ENMs were separated (Figure 5.1a). The SSD of ungrouped Ag ENMs is also enclosed for comparison. As can be observed from this figure, the SSD of the PVP-coated Ag ENMs shifted to the left compared with that of the uncoated Ag ENMs, which means that a PVP coating may considerably enhance the toxicity of Ag ENMs to most species. This agrees with the results obtained by Garner et al. (2015). Similarly, the sodium citrate-coated Ag ENMs also showed increased toxicity at high concentrations compared with the uncoated ones. As reported, both PVP and citrate are able to significantly reduce the aggregation and deposition to surfaces, and thus increase the bioavailability and toxicity (Gutierrez et al., 2015). The SSD of ungrouped Ag ENMs showed little statistical difference from that of the uncoated Ag ENMs. This could possibly be due to the counteraction of the influences of all kinds of surface coatings on the toxicity of Ag ENMs. The estimated HC5 value of uncoated Ag ENMs is 0.0063 mg/L, with the 95% CI ranging from 0.00098 to 0.068 mg/L. The HC5 of ungrouped Ag ENMs is 0.0036 mg/L (0.00064-0.029 mg/L). The HC5 of PVP-coated Ag ENMs is 0.0011 mg/L (0.00012-0.031 mg/L), and that of the sodium citrated-coated Ag ENMs is 0.0030 mg/L (0.00040-0.050 mg/L).

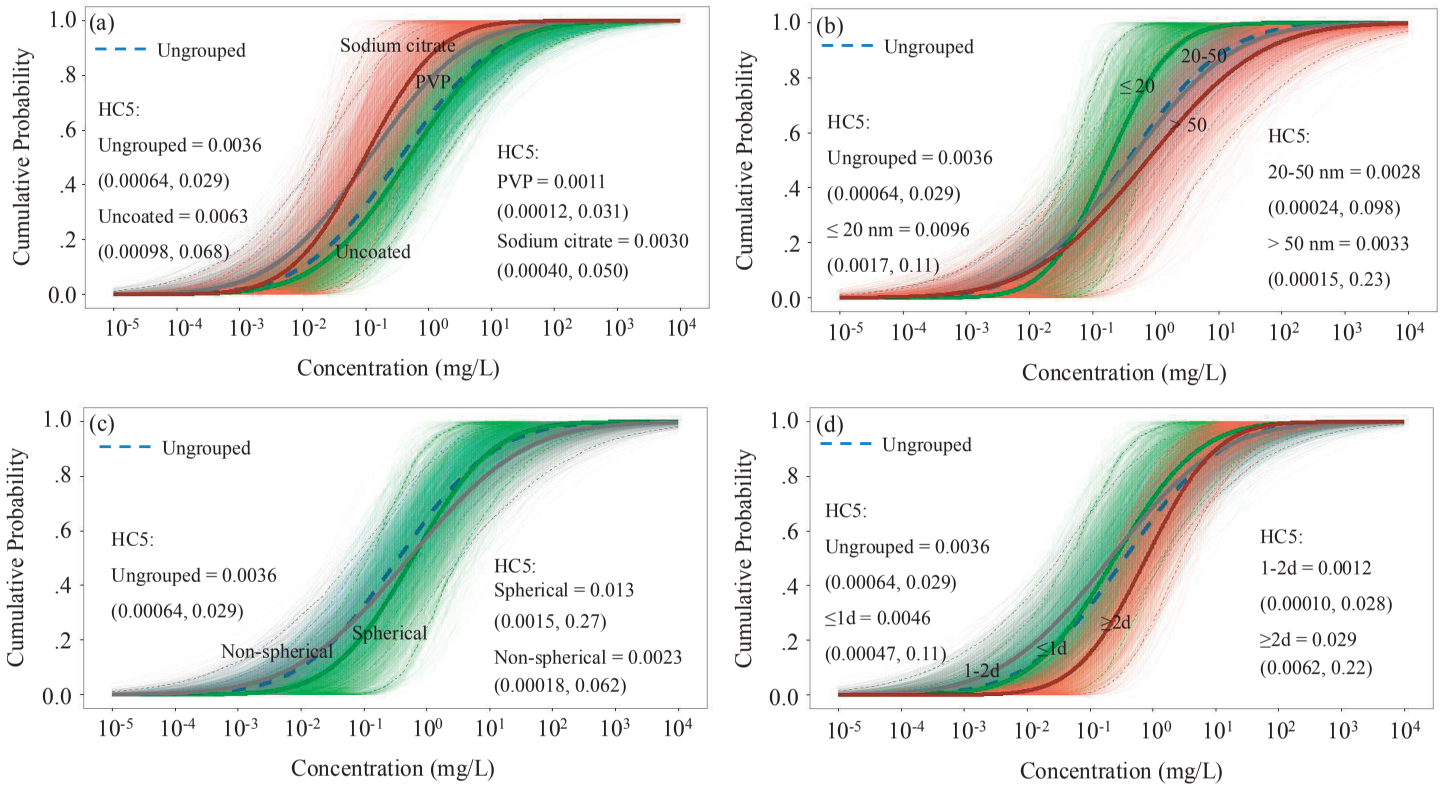


Figure 5.1. SSDs of Ag ENMs distinguished by (a) surface coating; (b) size; (c) shape; and (d) duration of toxicity exposure, based on LC50 values. The SSD of ungrouped Ag ENMs was also depicted in each figure for the comparison (dashed blue line). The shaded region along each curve shows the 95% confidence interval. The Ag ENMs in figures (b), (c) were not reported as being surface coated.

Grouped according to different size clusters of 1-20, 20-50, and 50-100 nm, the data were also ranked to create SSDs for Ag ENMs of different sizes (no surface coating reported), as shown in Figure 5.1b. Only minor differences were seen between the three SSDs especially at low concentrations, even though ENMs with smaller sizes are expected to act differently (Auffan et al., 2009; 2010). The SSD of ungrouped Ag ENMs unsurprisingly lies between those of Ag ENMs of 1-20 and 50-100 nm, which is nearly identical to the SSD of Ag ENMs with sizes ranging from 20 to 50 nm. The difference in behavior and/or effect is seen according to the separate SSDs when the exposure concentration increases; the group of smallest Ag ENMs tends to be relatively more toxic compared with the other two groups. One possible explanation for this observation is that the biological effects triggered by Ag ENMs are most likely to result from the release of Ag^+ ions (Juling et al., 2016). Therefore regardless of sizes, the mode of action of Ag ENMs of different sizes at low concentrations may be similar. As concentration rises, the proportion of the particle form significantly increases and ENM characteristics like size may start to play a role in affecting the toxicity. The study of Xiao et al. (2015) showed that the relative contribution of the particle forms of Cu ENMs to the accumulation in *Daphnia magna* increased from 48% to 72% when the concentrations of ENM suspensions increased from 0.05 to 0.1 mg/L. The same applies for the ZnO ENMs, as the relative contribution of their particle forms increased with the rise of concentrations of ZnO ENM suspensions (from 47% to 64% as concentration rised from 0.5 to 1 mg/L). The HC5 value of Ag ENMs ranging from 1 to 20 nm is 0.0096 mg/L (0.0017-0.11 mg/L). For Ag ENMs of 20-50 and 50-100 nm, the established HC5s are 0.0028 mg/L (0.00024-0.098 mg/L) and 0.0033 mg/L (0.00015-0.23 mg/L), respectively. SSDs of ENMs with different ranges of sizes were also derived for CuO, TiO_2, and ZnO ENMs as shown in Figure S5.2. The SSDs of CuO and TiO_2 ENMs distinguished by size highly overlap with those of the corresponding ungrouped ENMs within 95% CI. The SSD developed for ZnO ENMs of 50-100 nm also overlaps with that of the ungrouped ZnO ENMs especially at low concentrations.

Grouped within different shapes of ENMs, the data was also ranked to create SSDs. On the basis of the available data, only for spherical-shaped Ag ENMs (no reported coatings) a sufficient number of data points is available for the modeling. We therefore grouped the Ag ENMs as spherical and non-spherical Ag ENMs to determine if there are major differences between the distributions, as shown in Figure 5.1c. A comparison shows that the SSDs for spherical- and non-spherical- shaped Ag ENMs are nearly identical and the differences are minimal within corresponding 95% CI. Also the 95% CI of the ungrouped Ag ENMs heavily overlaps with those of the ENMs grouped by shape. This similarity could be possibly caused by the physical-chemical transformations of the particles in the medium, of which aggregation, agglomeration, and dissolution are the most important processes that alter the behaviors of ENMs and thereby the interactions of ENMs with biota (Chen et al.,

2015; Hua et al., 2016). In this context, the shape of Ag ENMs seems to play a less important role in influencing the toxicity of the materials. The calculated HC5 of non-spherical Ag ENMs is 0.0023 mg/L with the 95% CI ranging from 0.00018 to 0.062 mg/L. The HC5 value of the SSD of spherical Ag ENMs is equal to 0.013 mg/L (0.0015-0.27 mg/L).

The exposure duration used in the toxicity testing (Figure 5.1d) and light exposure (Figure S5.2) were also considered when constructing SSDs for metallic ENMs. No major statistical differences were seen between the ungrouped SSDs and the SSDs with distinct groups of species ranked as being exposed for ≤1 d and 1-2 d. Even so, at high concentrations (particularly above 10 mg/L) the three distributions highly overlap. HC5s derived from the SSDs of exposure duration ≤1 d and in between 1-2 d are 0.0046 mg/L (0.00047-0.11 mg/L) and 0.0012 mg/L (0.00010-0.028 mg/L), respectively. The HC5 generated from SSD of ≥ 2 d is 0.029 mg/L (0.0062-0.22 mg/L). For the toxicity of ENMs under different light exposures, most experiments followed standardized protocols such as OECD 202 (OECD, 2004) and US EPA (US EPA, 2002) which recommend a 16/8 h-light/dark-cycle for the toxicity testing. However, different lighting regimes were found to be applied for the toxicity test of TiO_2 ENMs due to their photoactivated toxicity. Sufficient data points based on EC50 (six species) were obtained only for UV exposed TiO_2 ENMs, and these were used in building the relevant SSD together with the SSD of ungrouped TiO_2 ENMs based on EC50 (Figure S5.2). As can be observed from the figure, the 95% CI of the SSD for UV exposed TiO_2 ENMs is much wider given the much smaller number of data points, which almost fully covers the 95% CI of the SSD for ungrouped TiO_2 ENMs (16 species). The HC5 value with respect to the ungrouped TiO_2 ENMs based on EC50 is 0.57 mg/L (0.16-2.8 mg/L), the HC5 estimated for the UV exposed TiO_2 ENMs is 1.5 mg/L with a 95% CI of 0.24-21 mg/L.

For the purpose of environmental risk assessment of ENMs, the variation of the obtained HC5s with 95% CI is of interest, as depicted in Figure 5.2 for Ag ENMs. As observed, most of the values of HC5s fall within the range of 10^{-3} to 10^{-2} mg/L with established 95% CI mainly ranging from 10^{-4} to 10^{-1} mg/L. Almost all the calculated 95% CIs of the HC5s highly overlap. This indicates that there are actually no statistically significant differences between the estimated HC5s from the SSDs of grouped or ungrouped Ag ENMs. The obtained HC5s of CuO, TiO_2, and ZnO ENMs were also depicted in Figure S5.3. Also no statistically significant differences were observed between the HC5s of relevant grouped and ungrouped ENMs.

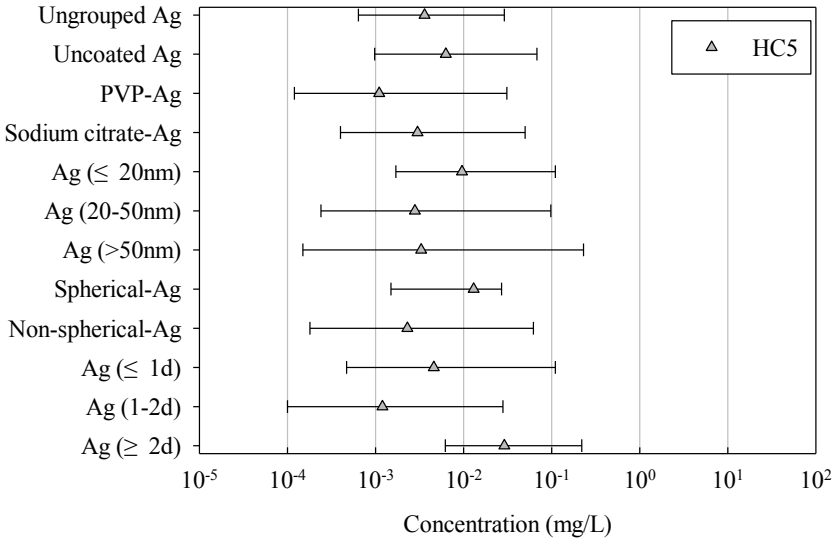


Figure 5.2. Comparison of HC5 values derived from SSDs of Ag ENMs differentiated by surface coating, size, shape, and exposure duration. Error bars show the 95% confidence interval of HC5s.

5.3.3 SSDs based on different toxicity endpoints

To compare the SSDs of certain ENMs based on different endpoints, the fitted distributions in Figure S5.1 were reorganized according to the type of ENM (Figure 5.3). Unexpectedly, only the SSDs of TiO_2 ENMs exhibited a reasonable order of NOEC < EC50 < LC50 at low concentrations. For Ag ENMs the difference is minimal between the NOEC- and LOEC-based SSDs, and also between the LC50- and EC50- based SSDs when the concentration is low. As concentration rises an order of NOEC < LOEC < LC50 < EC50 is seen. In the case of ZnO ENMs, major differences only appeared between the NOEC-SSDs and the SSDs based on other endpoints. The SSDs of ZnO ENMs based on EC50, LOEC, and NOEC showed no significant difference. This also applied for the NOEC- and LC50- SSDs, and the LOEC- and EC50- SSDs of CuO ENMs. The HC5s derived from these SSDs were calculated and compared in Figure 5.4. Based on LC50 (also see Figure S5.1), HC5s of the ENMs in an ascending order is Ag (0.0036 mg/L) < ZnO (0.022 mg/L) < CuO (0.049 mg/L) < TiO_2 (3.1 mg/L); For the HC5s based on EC50, it is Ag (0.0057 mg/L) < ZnO (0.058 mg/L) < CeO_2 (0.16 mg/L) < TiO_2 (0.57 mg/L) < CuO (1.3 mg/L); the order of LOEC-HC5s is Ag (0.00018 mg/L) < ZnO (0.086 mg/L) < CuO (3.2 mg/L); and in the case of NOEC the order is Ag (0.00036 mg/L) < ZnO (0.0051 mg/L) < CeO_2 (0.057 mg/L) < CuO (0.087 mg/L) < TiO_2 (0.19 mg/L).

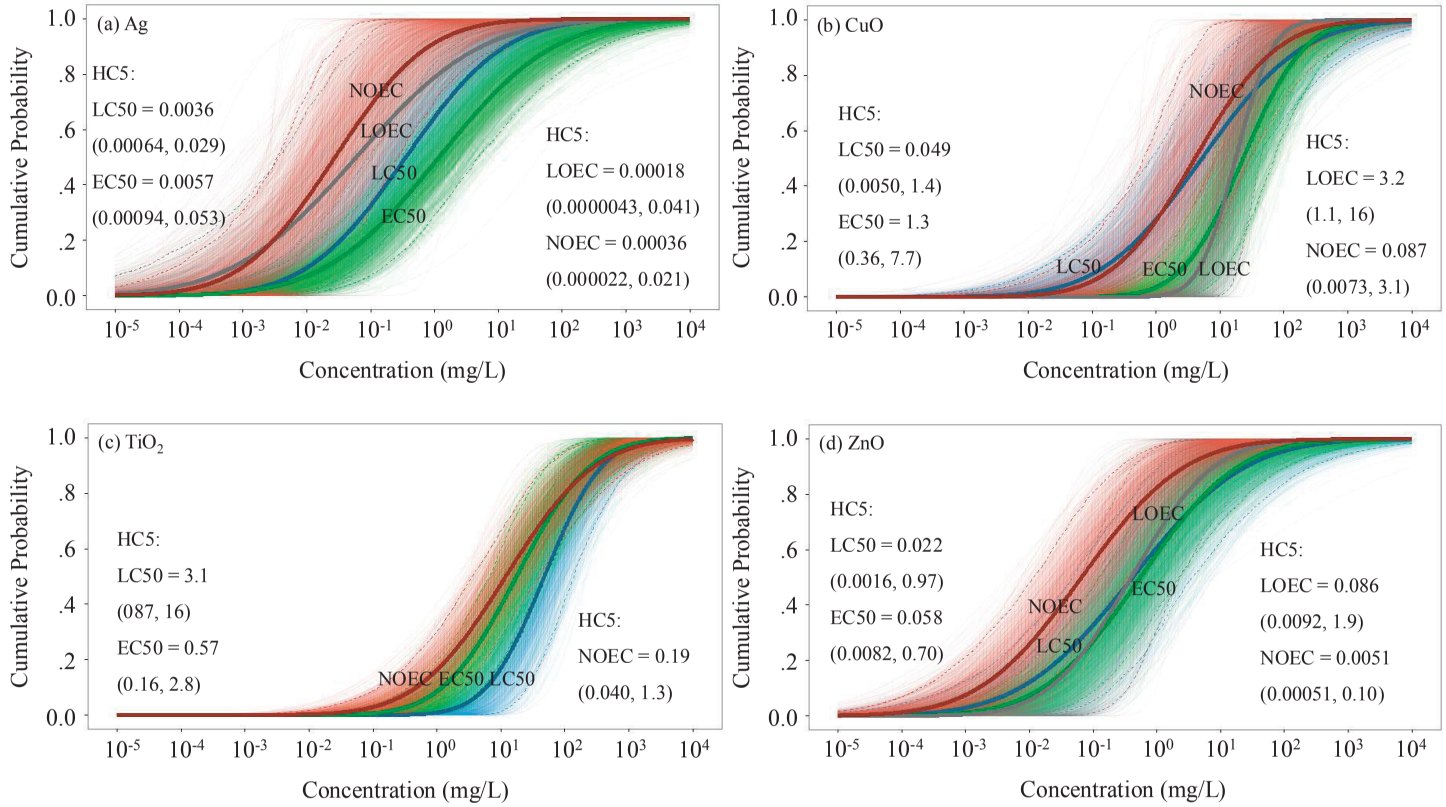


Figure 5.3. SSDs of (a) Ag; (b) CuO; (c) TiO_2; and (d) ZnO ENMs based on respectively LC50, EC50, LOEC, and NOEC data. The shaded region along each curve depicts the 95% confidence interval.

Interestingly, in all cases the HC5s of Ag and ZnO ENMs were shown to be lower than those of the other ENMs considered, whereas the ranking of the toxicity of CuO and TiO_2 ENMs differs when considering different toxicity endpoints. The predicted HC5s of Ag ENMs are always the lowest, and the toxicity of TiO_2 ENMs is commonly the lowest as can be concluded from the HC5 values (Garner et al., 2015; Coll et al., 2016; Gottschalk et al., 2013). As can be seen from Figure 5.4, the 95% CI of Ag ENMs is clearly significantly different from that of TiO_2 ENMs with no overlap of 95% CI with respect to any endpoint considered. This situation changes for the 95% CI of Ag and CuO ENMs which appear to be significantly different on the basis of EC50 and LOEC, but overlap when LC50 and NOEC are used. The conclusions of comparing HC5s (with 95% CI) of different ENMs vary when different endpoints are employed for modeling SSDs. For each ENM, no significant difference was found when comparing the NOEC-based HC5s with the HC5 values based on LC50, EC50, and LOEC, even though the NOEC-HC5s tend to be the lowest as concluded from the cases of CeO_2, TiO_2, and ZnO ENMs. Additionally, the ratios of LC50-HC5/NOEC-HC5, EC50-HC5/NOEC-HC5, and LOEC-HC5/NOEC-HC5 were calculated as listed in the Table S5.1. The ratio of LC50-HC5/NOEC-HC5 ranges from 0.6 (CuO ENMs) to 16.3 (TiO_2 ENMs). The ratio of EC50-HC5/NOEC-HC5 was found to range from 2.8 (CeO_2 ENMs) to 15.8 (Ag ENMs). With respect to LOEC-HC5/NOEC-HC5, the values vary from 0.5 (Ag ENMs) to 36.8 (CuO ENMs).

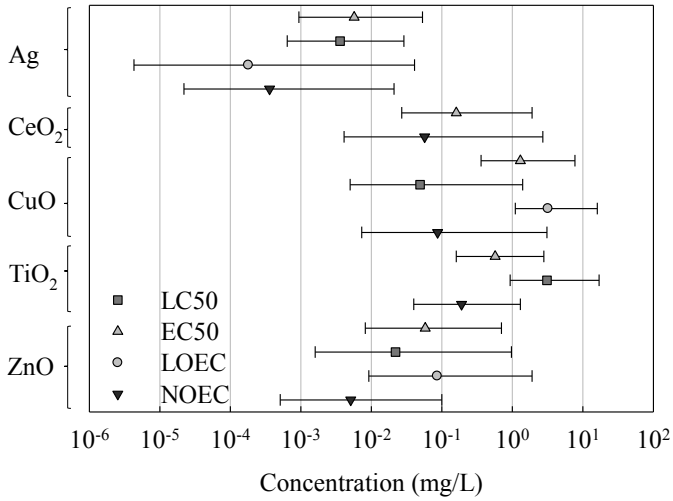


Figure 5.4. Variation of HC5 values of Ag, CeO_2, CuO, TiO_2, and ZnO ENMs based on respectively LC50, EC50, LOEC, and NOEC data. The 95% confidence interval is also given as well as the HC5 values.

5.4 Discussion

5.4.1 Data availability

Even though a large dataset (more than 1800 records) has been retrieved from >300 publications or scientific reports, it seems like so far only a limited number of ENMs were thoroughly investigated with regard to their toxicity to only a limited number of test species (Chen et al., 2015). When developing SSDs for the grouped ENMs, the data availability becomes even scarcer because of the lack of the data on, for example, ENM surface coatings, sizes, shapes, experimental conditions, etc. which are crucial for distinguishing the ENMs. The absence of these data could be due to the lack of data in original articles, or the missing of data when extracting information from publications to databases. In the present study, SSDs could only be developed for Ag, CeO_2, CuO, TiO_2, and ZnO ENMs based on all possible endpoints. According to the study of Bondarenko et al. (2013), Ag, CeO_2, CuO, TiO_2, and ZnO ENMs are indeed among the ENMs that are produced at the highest amounts, together with AlO_x, FeO_x, and SiO_2 ENMs. It would benefit the risk assessment of ENMs if all these metallic nanomaterials that are produced in high amounts were comprehensively evaluated for their safety, as they are all considered to inevitably enter into the environment and potentially pose impacts on human beings and environmental species (Echegoyen and Nerín, 2013). Developing SSDs for those ENMs of concern is one of the keys to manage the risks brought by the marketed nanomaterials. This nevertheless requires more types of ENMs to be tested, and also more relevant reliable models to be developed to reduce the time consumption and accelerate the process of risk evaluation. For the previously studied ENMs, toxicity data covering a wider range of taxa and trophic levels other than only standard species are also of significant importance to minimize the variabilities and levels of uncertainties.

In part, the data availability in developing SSDs also depends on firstly if the experimental results derived from a wide variety of protocols should be combined for building one SSD; and secondly, on the required minimum number of data points (number of species) to generate an SSD. Ideally, a distribution of species sensitivity ought to be generated from experiments that employed consistent protocols, for example, by using toxicity data reflecting the inhibition of growth or reproduction, or mortality, etc (Garner et al., 2015). In this context, only experimental results reflecting exactly the same biological effects should be grouped and used for the development of SSDs. This unquestionably largely reduces the available data for the modeling. According to the standardized toxicity testing protocols, different effects are recommended to be assessed for different standard test species, e.g., growth inhibition for *Pseudokirchneriella subcapitata* (OECD 201), immobility (OECD 202) and reproduction inhibition (OECD 211) for *Daphnia magna*, lethality for *Oryzias latipes*

(OECD 203) and *Danio rerio* embryo (OECD 236), etc. (OECD, 1992, 2004, 2011, 2012, 2013). Given the scarcity of data, it is as yet technically infeasible to include most of the species tested so far in one single SSD on the basis of one consistently measured effect level other than lethality. Therefore, data manipulation was adapted in previous studies so as to combine data representing different biological effects and to perform regression analysis (Coll et al., 2016; Garner et al., 2015; Gottschalk et al., 2013). Additionally, the minimum number of data points to build an SSD also determines whether a dataset with a very limited number of species can be used for modeling. Although it was proposed by Garner et al. (2015) that a minimum of four species is needed to construct SSDs, Cedergreen et al. (2004) stated that at least six to eight species must be represented. Therefore, assuming that only four data points are required for the SSD derivation, the SSDs for Cu and Ni ENMs could also be built based on LC50 data (Table 5.1). This will however induce a quite broad CI.

5.4.2 Comparison of SSDs and relevant HC5s

Given the relatively high amount of data, SSDs could be built for Ag ENMs distinguished by coating, range of size, shape, and exposure duration. Although a few of the distributions (e.g., SSDs in Figure 5.1b) at high concentrations showed some variations, the HC5s that were derived from the developed SSDs do not differ significantly. This means that, on the basis of the currently available data, all kinds of Ag ENMs entering into the environment are supposed to share similar maximum acceptable concentrations, regardless of surface coatings, shapes, sizes, exposure durations, or even other structural characteristics. This similarity could possibly result from either or both of the two major reasons. The first is the physical-chemical transformation of Ag ENMs in the aquatic media which can completely change the structural properties of ENMs (Chen et al., 2015). Despite the fact that the structural parameters of ENMs have been formerly linked to the toxicity of ENMs (Chen et al., 2016), it is still difficult to quantify the relationship between the characteristics of pristine ENMs (e.g., size, surface coating, shape, etc.) and the behaviors of ENMs in a medium. This behavior may alter the mobility, bioavailability, and ultimately the toxicity of the nanomaterials, and thus is of vital significance to understand the mechanisms governing nanotoxicity. The second reason is the general mechanism of toxicity of nano-, micro-, and bulk- Ag releasing metal ions. As known, one of the major mechanisms of Ag-induced toxicity is the leaching of Ag^+ ions. Therefore especially at low concentrations, Ag ENMs with varied structural properties tend to exhibit analogous biological activities. But as concentrations increase, the proportion of the nanoparticulate Ag will as well rise and differences would probably emerge between the SSDs of Ag ENMs with different structural properties. As for the influence of light exposure, the SSD could only be developed for the UV exposed TiO_2 ENMs which is incomparable. The different is not significant either

between the SSDs of UV exposed and ungrouped TiO_2 ENMs based on EC50 (Figure S5.2).

Assessment factors are commonly used when deriving the predicted no observed effect concentrations from the HC5s. For instance, a factor of 10 was used by Gottschalk et al. (2013) to calculate the predicted no observed effect concentrations from LC50 and EC50, while a factor of 2 was applied to generate this value from LOEC. In the study of Coll et al. (2016), a factor of 10 was used for LC50 and EC50, and a value of 1 was employed for LOEC and NOEC. Based on our results, the ratio of HC5s of L(E)C50/NOEC ranges from 0.6 to the highest 16.3 with a median value of 10 (Table S5.1). For the combination of LOEC/NOEC the three values are 0.5 (Ag ENMs), 16.9 (ZnO ENMs), and 36.8 (CuO ENMs). Although the limited number of data points of Ag and CuO ENMs (only five data points for both LOEC and NOEC data, see Table 5.1) will cause larger uncertainties, the value of 16.9 (LOEC-HC5/NOEC-HC5) for ZnO ENMs with a relatively sufficient number of data (respectively 6 and 11 for LOEC and NOEC data) does not seem to be close to a factor of 2. With respect to the SSDs built on different toxicity endpoints, the NOEC-SSDs were not as expected significantly lower than that based on LC50, EC50, and LOEC except for the case of ZnO ENMs. Neither did the LOEC-SSDs always appear in between the NOEC-SSDs and the EC50-SSDs, as expected on forehand. Given the situation that NOEC should always represent the most sensitive case, the ratio of L(E)C50-HC5/NOEC-HC5 and LOEC-HC5/NOEC-HC5 was actually also not considered to be lower than 1. This was however observed for the LOEC-HC5/NOEC-HC5 of Ag ENMs (0.5) and for the LC50-HC5/NOEC-HC5 of CuO ENMs (0.6). Together with the discussed discrepancies of SSDs in Figure 5.3, we understand that this might be attributed to the fact that the data used were retrieved from a variety of sources with varying data quality. The limited sample sizes of Ag and CuO ENMs based on respectively LOEC and NOEC also resulted in the wide CI and low statistical power. These uncertainties could only be diminished by future increase of data quality and availability.

5.4.3 Most sensitive species and organism groups

Based on the developed SSDs, we listed the most sensitive species of every SSD in Table S5.2. Despite that no single species was found to be always the most susceptible, a few species were constantly observed to be the most vulnerable to metallic ENMs. These species include *Ceriodaphnia affinis*, *Ceriodaphnia dubia*, *Daphnia magna*, *Daphnia pulex*, *Escherichia coli*, and *Pseudokirchneriella subcapitata*. Most of these species are crustaceans which account for 26 out of 32 of the most sensitive species in the SSDs developed. This indicates that crustaceans are more likely to be the organism group that is affected by the metal-based ENMs at the lowest concentrations of ENMs. This observation is in line with the study of

Garner et al. (2015), in which the most sensitive species to metallic ENMs were all crustaceans, namely *Ceriodaphnia dubia* (in SSDs of uncoated and PVP-coated Ag, Al_2O_3, Cu, and TiO_2 ENMs), *Daphnia pulex* (CuO ENMs), *Daphnia similis* (CeO_2 ENMs), and *Thamnocephalus platyurus* (ZnO ENMs). Since the HC5 represents a concentration where only 5% of the species could be affected, it seems that the crustaceans would be those that are within the 5% of the species. Therefore in the case of a generic risk assessment, it may be important to include at least a few representative species from the crustacean group in the SSDs such as *Ceriodaphnia dubia*, *Daphnia magna*, and *Daphnia pulex*.

5.4.4 Conclusions

To conclude, reliable information on the characteristics of ENMs that govern toxicity and the experimental conditions are needed for the development of separate SSDs. More data on the highly produced ENMs such as AlO_x, CeO_2, CuO, FeO_x, SiO_2, TiO_2, and ZnO ENMs are favorable for a comprehensive evaluation of the environmental risks of ENMs. Sufficient data on Ag ENMs enabled a comparison between the SSDs constructed for the grouped Ag ENMs. For the Ag ENMs grouped by shape and exposure duration, the separate SSDs of Ag ENMs showed no statistically significant difference. For the Ag ENMs of different size ranges, differences in behavior and/or effect were only seen at high exposure concentrations. The PVP- and sodium citrate- coatings on the surface of Ag ENMs enhance the nanotoxicity as the SSDs shifted to the left compared to the SSD of the uncoated Ag ENMs. The derived HC5s for all the grouped Ag ENMs do not differ significantly, which implies that only the intrinsic chemical toxicity of Ag ENMs greatly affected the corresponding SSDs. HC5s generated from the SSDs of ungrouped Ag, CeO_2, CuO, TiO_2, and ZnO ENMs based on respectively LC50, EC50, LOEC, and NOEC were also compared. Median values of 10 for the ratio of L(E)C50-HC5/NOEC-HC5, and of 16.9 for the ratio of LOEC-HC5/NOEC-HC5 were obtained. An analysis of the most sensitive species in every SSD showed that no single species was consistently the most sensitive, however crustaceans as an organism group tend to be extra vulnerable to metal-based ENMs. Due to the limitations caused by data quality and availability, it should be noticed that uncertainties still exist associated with our results. For the developed SSDs, such uncertainties could be reduced if reliable toxicity information of sufficient species became available which could represent a comprehensive ecosystem. Despite these considerations, we believe the present study is helpful in gauging the SSDs of ENMs grouped by individual ENM properties and other important factors, and in enabling the further development of SSDs for metallic ENMs.

Acknowledgements

Special thanks to E. Szöcs (University of Koblenz-Landau) for providing guidance in using R software. G. Chen greatly thanks the funding support by the Chinese Scholarship Council (201306060076). M.G. Vijver is funded by the NWO-VIDI 864.13.010. Part of the work was performed within the framework of the EU-sponsored project "FutureNanoNeeds", grant agreement number 604,602.

References

Auffan, M., Bottero, J.Y., Chaneac, C., Rose, J., 2010. Inorganic manufactured nanoparticles: how their physicochemical properties influence their biological effects in aqueous environments. Nanomedicine 5, 999-1007.

Auffan, M., Rose, J., Bottero, J.Y., Lowry, G.V., Jolivet, J.P., Wiesner, M.R., 2009. Towards a definition of inorganic nanoparticles from an environmental, health and safety perspective. Nat. Nanotechnol. 4, 634-41.

Bondarenko, O., Juganson, K., Ivask, A., Kasemets, K., Mortimer, M., Kahru, A., 2013. Toxicity of Ag, CuO and ZnO nanoparticles to selected environmentally relevant test organisms and mammalian cells in vitro: a critical review. Arch. Toxicol. 87, 1181-200.

Cedergreen, N., Spliid, N.H., Streibig, J.C., 2004. Species-specific sensitivity of aquatic macrophytes towards two herbicide. Ecotoxicol. Environ. Saf. 58, 314-23.

Chen, G., Peijnenburg, W.J.G.M., Kovalishyn, V., Vijver, M.G., 2016. Development of nanostructure–activity relationships assisting the nanomaterial hazard categorization for risk assessment and regulatory decision-making. RSC Adv., 6, 52227-52235.

Chen, G., Vijver, M.G., Peijnenburg, W.J.G.M., 2015. Summary and analysis of the currently existing literature data on metal-based nanoparticles published for selected aquatic organisms: Applicability for toxicity prediction by (Q)SARs. Altern. Lab. Anim. 43, 221-40.

Coll, C., Notter, D., Gottschalk, F., Sun, T., Som, C., Nowack, B., 2016. Probabilistic environmental risk assessment of five nanomaterials (nano-TiO2, nano-Ag, nano-ZnO, CNT, and fullerenes). Nanotoxicology. 10, 436-44.

ECHA (European Chemicals Agency), 2008. Guidance on information requirements and chemical safety assessment. Helsinki, Finland.

Echegoyen, Y., Nerín, C., 2013. Nanoparticle release from nano-silver antimicrobial food containers. Food Chem. Toxicol. 62, 16-22.

Garner, K.L., Suh, S., Lenihan, H.S., Keller, A.A., 2015. Species sensitivity distributions for engineered nanomaterials. Environ. Sci. Technol. 49, 5753-9.

Gottschalk, F., Kost, E., Nowack, B., 2013. Engineered nanomaterials in water and soils: a risk quantification based on probabilistic exposure and effect modeling. Environ. Toxicol. Chem. 32, 1278-87.

Gottschalk, F., Nowack, B., 2013. A probabilistic method for species sensitivity distributions taking into account the inherent uncertainty and variability of effects to estimate environmental risk. Integr. Environ. Assess. Manag. 9, 79-86.

Gutierrez, L., Aubry, C., Cornejo, M., Croue, J.P., 2015. Citrate-Coated Silver Nanoparticles Interactions with Effluent Organic Matter: Influence of Capping Agent and Solution Conditions. Langmuir 31, 8865-72.

Hua, J., Vijver, M.G., Chen, G., Richardson, M.K., Peijnenburg, W.J.G.M., 2016. Dose metrics assessment for differently shaped and sized metal-based nanoparticles. Environ. Toxicol. Chem. 35, 2466-2473.

Hyndman, R. J., Fan, Y., 1996. Sample quantiles in statistical packages. Am. Stat. 50, 361-365.

Iversen, T.G., Skotland, T., Sandvig, K., 2011. Endocytosis and intracellular transport of nanoparticles: Present knowledge and need for future studies. Nano Today 6, 176-185.

Jin, H., Heller, D.A., Sharma, R., Strano, M.S., 2009. Size-dependent cellular uptake and expulsion of single-walled carbon nanotubes: Single particle tracking and a generic uptake model for nanoparticles. ACS Nano 3, 149-158.

Juganson, K., Ivask, A., Blinova, I., Mortimer, M., Kahru, A., 2015. NanoE-Tox: New and in-depth database concerning ecotoxicity of nanomaterials. Beilstein J. Nanotechnol. 6, 1788-804.

Juling, S., Bachler, G., von Götz, N., Lichtenstein, D., Böhmert, L., Niedzwiecka, A., Selve, S., Braeuning, A., Lampen, A., 2016. In vivo distribution of nanosilver in the rat: The role of ions and de novo-formed secondary particles. Food Chem. Toxicol. 97, 327-335.

OECD (Organization for Economic Co-operation and Development), 1992. OECD Guideline for Testing of Chemicals. No. 203, Fish, Acute Toxicity Test.

OECD (Organization for Economic Co-operation and Development), 2004. OECD Guideline for Testing of Chemicals. No. 202, Daphnia sp., Acute Immobilization Test.

OECD (Organization for Economic Co-operation and Development), 2011. OECD Guideline for Testing of Chemicals. No. 201, Freshwater Alga and Cyanobacteria, Growth Inhibition Test.

OECD (Organization for Economic Co-operation and Development), 2012. OECD Guideline for Testing of Chemicals. No. 211, Daphnia magna Reproduction Test.

OECD (Organization for Economic Co-operation and Development), 2013. OECD Guideline for Testing of Chemicals. No. 236, Fish Embryo Acute Toxicity (FET) Test.

Piperigkou, Z., Karamanou, K., Engin, A.B., Gialeli, C., Docea, A.O., Vynios, D.H., Pavão, M.S., Golokhvast, K.S., Shtilman, M.I., Argiris, A., Shishatskaya, E., Tsatsakis, A.M., 2016. Emerging aspects of nanotoxicology in health and disease: From agriculture and food sector to cancer therapeutics. Food Chem. Toxicol. 91, 42-57.

Sanders, K., Degn, L.L., Mundy, W.R., Zucker, R.M., Dreher, K., Zhao, B., Roberts, J.E., Boyes, W.K., 2012. In vitro phototoxicity and hazard identification of nano-scale titanium dioxide. Toxicol. Appl. Pharmacol. 258, 226-36.

Sushko, I., Novotarskyi, S., Körner, R., Pandey, A.K., Rupp, M., Teetz, W., Brandmaier, S., Abdelaziz, A., Prokopenko, V.V., Tanchuk, V.Y., Todeschini, R., Varnek, A., Marcou, G., Ertl, P., Potemkin, V., Grishina, M., Gasteiger, J., Schwab, C., Baskin, I.I., Palyulin, V.A., Radchenko, E.V., Welsh, W.J., Kholodovych, V., Chekmarev, D., Cherkasov, A., Aires-de-Sousa, J., Zhang, Q.Y., Bender, A., Nigsch, F., Patiny, L., Williams, A., Tkachenko, V., Tetko, I.V., 2011. Online chemical modeling environment (OCHEM): web platform for data storage, model development and publishing of chemical information. J. Comput. Aided Mol. Des. 25, 533-54.

Toropova, A.P., Toropov, A.A., 2013. Optimal descriptor as a translator of eclectic information into the prediction of membrane damage by means of various TiO2 nanoparticles. Chemosphere. 93, 2650-2655.

US EPA (US Environmental Protection Agency), 1998. Guidelines for ecological risk assessment. EPA/630/R-95/002E Risk Assessment Forum. Washington, DC, USA.

US EPA (US Environmental Protection Agency), 2002. Methods for measuring the acute toxicity of effluents and receiving waters to freshwater and marine organisms. Fifth edition. EPA-821-R-02-012. Washington, DC: USEPA Office of Water.

Venables, W. N., Ripley, B. D., 2002. Modern Applied Statistics with S. Fourth edition. Springer.

Winkler, D.A., 2016. 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.

Xiao, Y., Vijver, M.G., Chen, G., Peijnenburg, W.J.G.M., 2015. Toxicity and accumulation of Cu and ZnO nanoparticles in Daphnia magna. Environ. Sci. Technol. 49, 4657-64.

Yin, J.J., Liu, J., Ehrenshaft, M., Roberts, J.E., Fu, P.P., Mason, R.P., Zhao, B., 2012. Phototoxicity of nano titanium dioxides in HaCaT keratinocytes—generation of reactive oxygen species and cell damage. Toxicol. Appl. Pharmacol. 263, 81-8.

Chapter 5 Supplemental Information

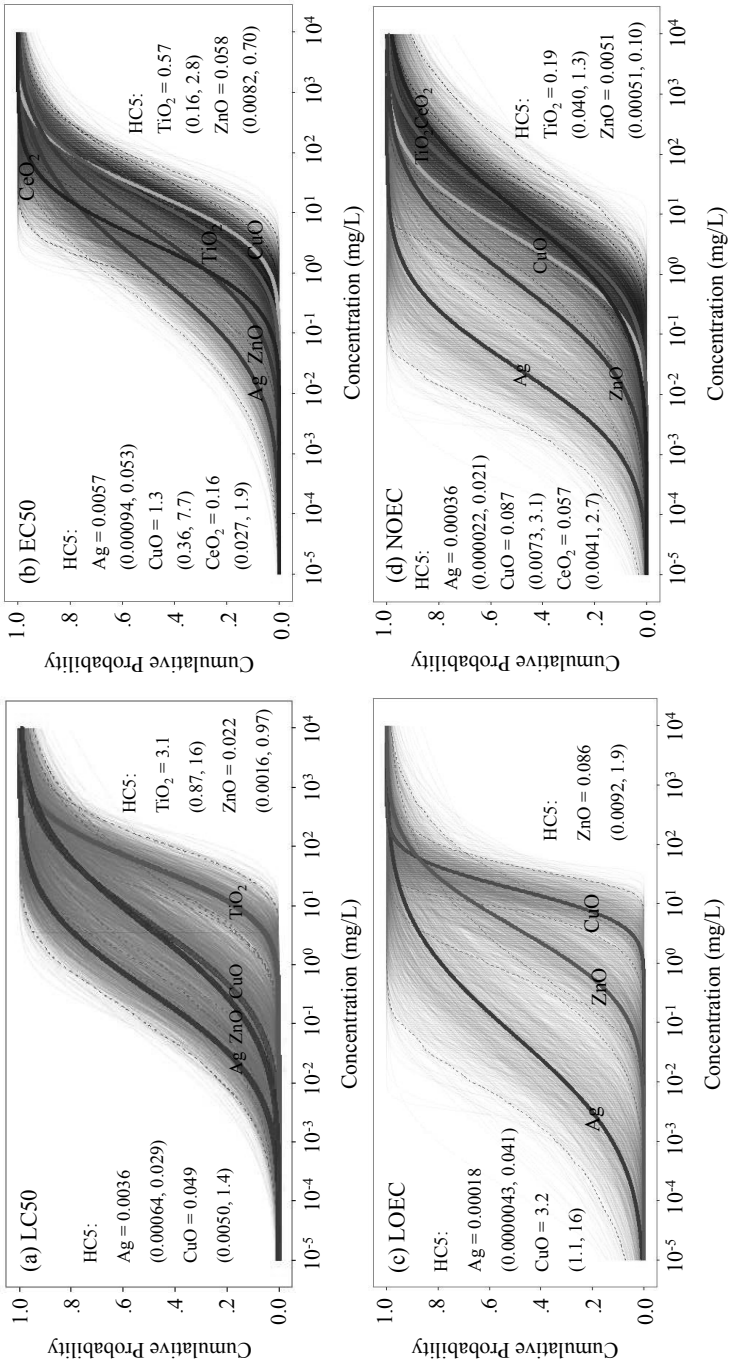


Figure S5.1. SSDs of Ag, CeO_2, CuO, TiO_2, and ZnO ENMs based on the toxicity endpoints of (a) LC50; (b) EC50; (c) LOEC; and (d) NOEC. The shaded region along each curve depicts the 95% confidence interval.

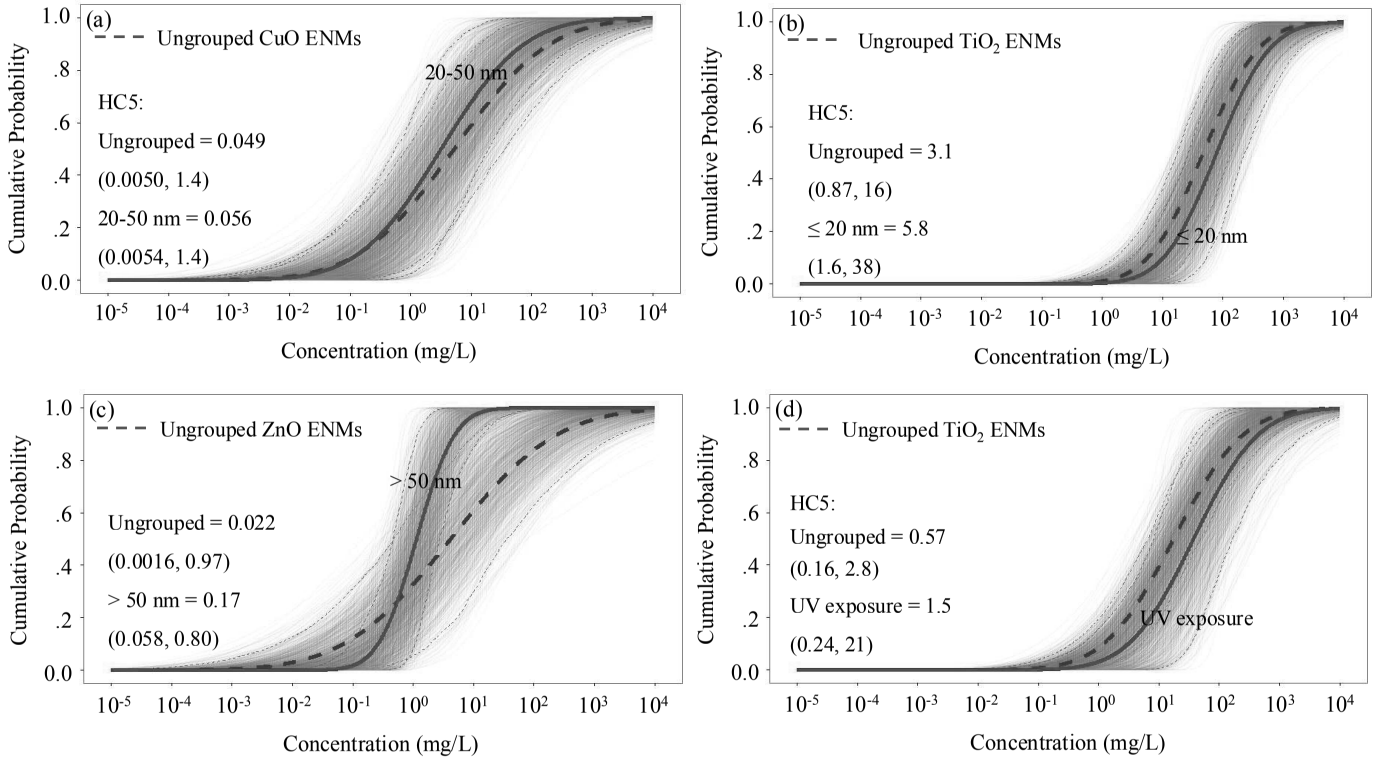


Figure S5.2. SSDs of (a) CuO; (b) TiO_2; and (c) ZnO ENMs with different size clusters based on LC50. The SSD of TiO_2 ENMs under UV exposure based on EC50 was depicted in (d). SSDs of unseparated ENMs were also enclosed for comparison (dashed line). The shaded region along each curve depicts the 95% confidence interval. All three ENMs grouped by size were not reported as being surface coated.

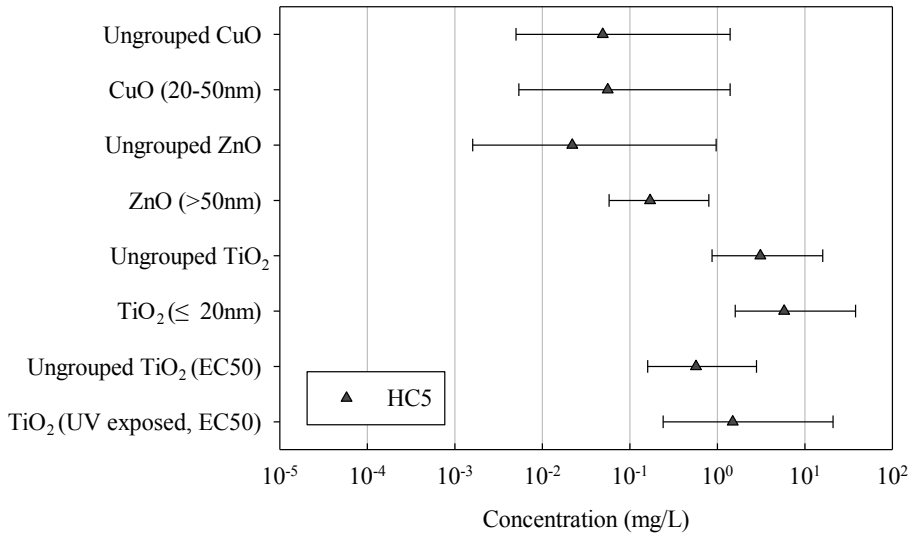


Figure S5.3. Comparison of HC5 values derived from SSDs of CuO, TiO_2, and ZnO ENMs separated by size (using LC50 data), and of TiO_2 ENMs separated by UV exposure (using EC50 data). Error bars show the 95% confidence interval of HC5s.

Table S5.1. HC5 values of LC50-, EC50-, and LOEC- based SSDs divided by the HC5 of NOEC-SSDs

	Ag	CeO_2	CuO	TiO_2	ZnO
LC50/NOEC	10.0		0.6	16.3	4.3
EC50/NOEC	15.8	2.8	14.9	3.0	11.4
LOEC/NOEC	0.5		36.8		16.9

Examples of developing SSDs for metallic ENMs

Example I: building SSDs for Ag ENMs separated by shape using LC50 data

Based on the retrieved LC50 data of Ag ENMs, the toxicity records of spherical Ag and non-spherical Ag ENMs were initially separated. Within each of the sub-dataset, the median toxicity value of Ag ENMs to each species was calculated, and ranked from the lowest to highest by the equation given in 5.2.2 Modeling algorithm of Chapter 5. The obtained values are as follows:

Ag Spherical

Species	LC50	Unit	Shape	Rank	Proportion
Daphnia magna	0.0175	mg/L	Spherical	1	0.0714286
Fathead minnow	0.0894	mg/L	Spherical	2	0.2142857
Pimephales promelas	0.09	mg/L	Spherical	3	0.3571429
Rainbow trout	0.71	mg/L	Spherical	4	0.5
Ceriodaphnia dubia	3.32	mg/L	Spherical	5	0.6428571
Moina macrocopa	5.77	mg/L	Spherical	6	0.7857143
Danio rerio	13.62483	mg/L	Spherical	7	0.9285714

Ag Non-spherical

Species	LC50	Unit	Shape	Rank	Proportion
Ceriodaphnia dubia	7.71E-04	mg/L	Non-spherical	1	0.0454545
Daphnia magna	0.00525	mg/L	Non-spherical	2	0.1363636
Daphnia pulex	0.04	mg/L	Non-spherical	3	0.2272727
Pseudokirchneriella subcapitata	0.19	mg/L	Non-spherical	4	0.3181818
Hypophthalmichthys molitrix	0.5155	mg/L	Non-spherical	5	0.4090909
Danio rerio	0.775	mg/L	Non-spherical	6	0.5
Oryzias latipes	1.03	mg/L	Non-spherical	7	0.5909091
Pimephales promelas	5.38	mg/L	Non-spherical	8	0.6818182
Oreochromis mossambicus	12.6	mg/L	Non-spherical	9	0.7727273
Rainbow trout	23.18	mg/L	Non-spherical	10	0.8636364
Paramecium caudatum	39	mg/L	Non-spherical	11	0.9545455

With these values, the lognormal distributions of species sensitivity were fitted using the 'fitdistr' function of the MASS package in the R statistical software (version 3.3.1), and HC5 values were also extracted by the 'quantile' function. The obtained SSDs are shown in Figure S5.4 (also in Figure 5.1c in Chapter 5). In the figure, data points reflecting the median values are shown together with the names of corresponding species. The shaded region of each curve depicts the 95% confidence interval (CI). The calculated HC5 of spherical Ag ENMs is 0.013 with CI ranging from 0.0015 to 0.27 mg/L, the HC5 of non-spherical Ag ENMs is 0.0023 (0.00018-0.062) mg/L.

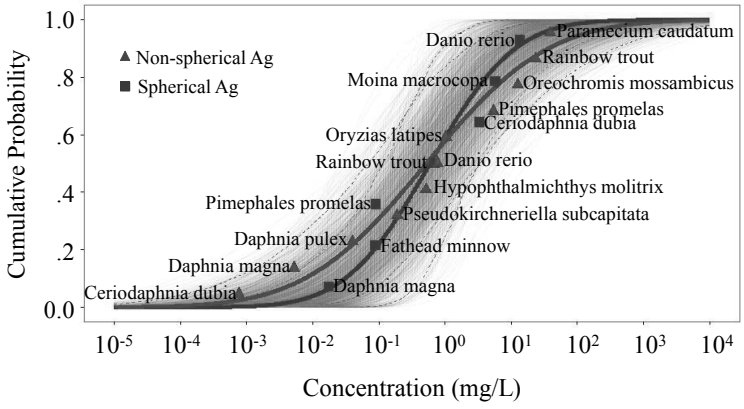


Figure S5.4. Developed SSDs for Ag ENMs separated by shape using LC50 data

Example II: building SSDs for ZnO ENMs using EC50 data

Different from LC50 data that is only based on mortality, the EC50 (and also LOEC and NOEC) dataset includes toxicity records on the basis of multiple biological effects. For example, in the EC50 data of ZnO ENMs, *Caenorhabditis elegans* was tested for both reproduction inhibition and immobilization; *Daphnia magna* was tested for feeding inhibition, reproduction inhibition, and immobilization. Thus to obtain only one toxicity value from the data of one species for plotting the distribution of species sensitivity, the median toxicity value based on a certain biological effect to a species was initially calculated per reported effect. In the case of *Caenorhabditis elegans* it is 790.67 mg/L (immobilization) and 57.3 mg/L (reproduction inhibition), for *Daphnia magna* it is 1.6685, 3.1, and 0.156 mg/L based on respectively feeding inhibition, immobilization, and reproduction inhibition. Afterwards, the obtained medians of different effects to that species were compared and the lowest median value was used in ranking the species sensitivities (e.g. 57.3 mg/L for *Caenorhabditis elegans*

based on reproduction inhibition; 0.156 mg/L for *Daphnia magna* based on reproduction inhibition). The ranked median values of different species were then plotted against the cumulative probability which reflects the proportion of species affected at a certain concentration. With the 'fitdistr' function of the MASS package in the R statistical software (version 3.3.1), the SSD for ZnO ENMs using EC50 data is obtained, as shown in Figure S5.5 (also in Figure 5.3d in Chapter 5, Fig. S5.1b). Data points with corresponding species names are also given in the figure.

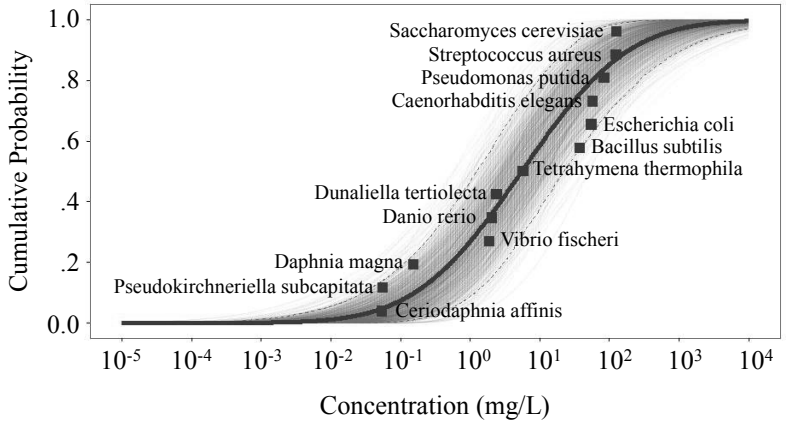


Figure S5.5. Developed SSD for ZnO ENMs using EC50 data

Example III & IV: developing SSDs for CuO (LOEC data) and TiO_2 (NOEC data) ENMs

Same as the data processing for building SSDs using EC50 data, when using LOEC and NOEC datasets, the median toxicity value based on a certain biological effect to a species was firstly obtained per reported effect, and then the lowest median value for a species was used in ranking the species sensitivities. The datasets used for building SSDs for CuO (LOEC data) and TiO_2 (NOEC data) ENMs were listed in the Supplemental Information (Microsoft Excel spreadsheet). The developed SSDs including information on species are shown in Figure S5.6 and Figure S5.7 as examples. These SSDs were also depicted in Figure 5.3 in Chapter 5 and in Figure S5.1 with simplified information for the purpose of conciseness.

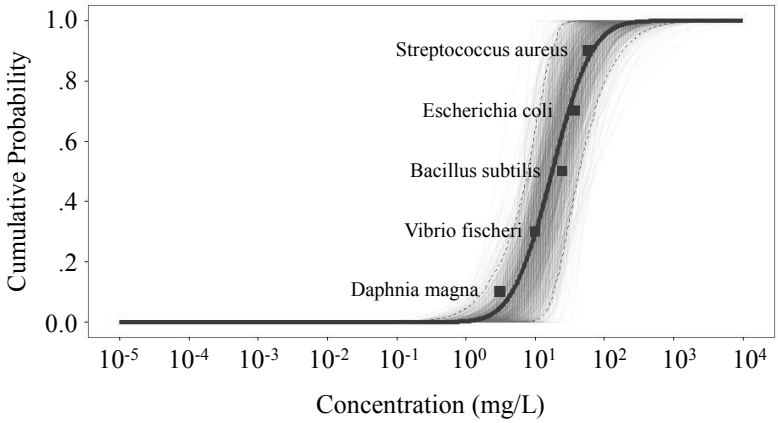


Figure S5.6. Developed SSD for CuO ENMs using LOEC data

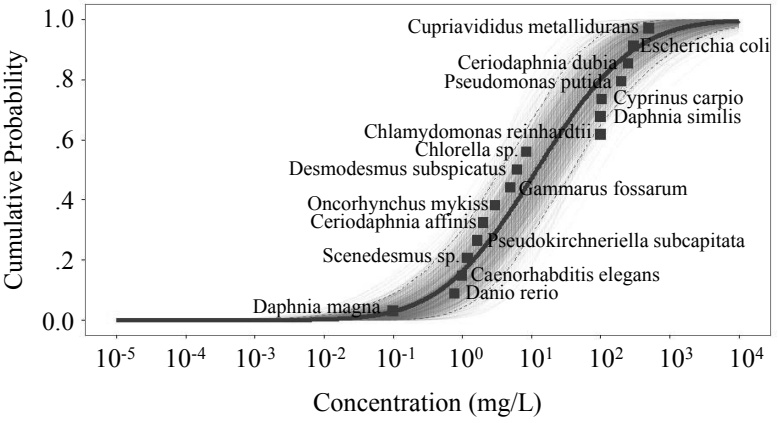


Figure S5.7. Developed SSD for TiO_2 ENMs using NOEC data

Table S5.2. Analysis of the most sensitive species in every developed SSD. Check mark means that some of the species presented in the SSD belong to the corresponding organism group

SSDs		Algae	Bacteria	Crustacean	Fish	Nematodes	Protozoa	Yeast	Most sensitive species	Most sensitive organism
LC50	Ungrouped Ag	✓		✓	✓	✓	✓		*Ceriodaphnia dubia*	crustacean
	Uncoated Ag	✓		✓	✓		✓		*Ceriodaphnia dubia*	crustacean
	PVP-coated Ag			✓	✓		✓		*Ceriodaphnia dubia*	crustacean
	Sodium citrate-coated Ag	✓		✓	✓				*Daphnia magna*	crustacean
	Ag ≤ 20 nm			✓	✓				*Daphnia magna*	crustacean
	Ag 20-50 nm	✓		✓	✓		✓		*Ceriodaphnia dubia*	crustacean
	Ag 50-100 nm			✓	✓				*Daphnia magna*	crustacean
	CuO 20-50 nm	✓	✓	✓					*Daphnia pulex*	crustacean
	TiO2 ≤ 20 nm		✓	✓	✓				*Daphnia pulex*	crustacean
	ZnO 50-100 nm		✓	✓	✓	✓			*Escherichia coli*	bacteria

Table S5.2. (Continued)

	SSDs	Algae	Bacteria	Crustacean	Fish	Nematodes	Protozoa	Yeast	Most sensitive species	Most sensitive organism
LC50	Ag Spherical			✓	✓				*Daphnia magna*	crustacean
	Ag Nonspherical	✓		✓	✓		✓		*Ceriodaphnia dubia*	crustacean
	Ag ≤ 1 d	✓		✓	✓		✓		*Daphnia magna*	crustacean
	Ag 1-2 d			✓	✓				*Ceriodaphnia dubia*	crustacean
	Ag ≥ 2 d	✓		✓	✓	✓			*Daphnia magna*	crustacean
	Ungrouped CuO	✓	✓	✓	✓		✓		*Daphnia pulex*	crustacean
	Ungrouped TiO_2		✓	✓	✓	✓			*Daphnia magna*	crustacean
	Ungrouped ZnO		✓	✓	✓	✓	✓		*Escherichia coli*	bacteria
EC50	Ag	✓	✓	✓	✓	✓	✓		*Daphnia magna*	crustacean
	CeO_2	✓	✓	✓					*Daphnia similis*	crustacean
	CuO	✓	✓	✓			✓	✓	*Pseudokirchneriella subcapitata*	algae

Table S5.2. (Continued)

	SSDs	Algae	Bacteria	Crustacean	Fish	Nematodes	Protozoa	Yeast	Most sensitive species	Most sensitive organism
EC50	TiO_2	✓	✓	✓	✓				*Ceriodaphnia affinis*	crustacean
	ZnO	✓	✓	✓	✓	✓	✓	✓	*Ceriodaphnia affinis*	crustacean
	UV exposed TiO_2	✓	✓	✓					*Pseudokirchneriella subcapitata*	algae
LOEC	Ag	✓	✓	✓	✓				*Daphnia magna*	crustacean
	CuO		✓	✓					*Daphnia magna*	crustacean
	ZnO	✓	✓	✓					*Acartia tonsa*	crustacean
NOEC	Ag	✓	✓	✓		✓			*Daphnia magna*	crustacean
	CeO_2	✓	✓	✓		✓			*Escherichia coli*	bacteria
	CuO	✓	✓	✓					*Thamnocephalus platyurus*	crustacean
	TiO_2	✓	✓	✓	✓	✓			*Daphnia magna*	crustacean
	ZnO	✓	✓	✓	✓				*Pseudokirchneriella subcapitata*	algae

CHAPTER 6

GENERAL DISCUSSION

This chapter was published with minor modification:

Chen G, Peijnenburg W, Xiao Y, Vijver MG

Published in *International Journal of Molecular Sciences*. 2017, 18(7): 1504

Nanotechnology has been identified as a key-enabling technology by the European Commission (European Commission, 2017). It is seen as one of the sectors bringing economic benefit and jobs. The extensive use of engineered nanomaterials (ENMs), however, has raised concerns about their possible effects on human health and their environmental burden (Nel et al., 2006). Laboratory observations on some potentially harmful effects of ENMs have in some cases overshadowed the immense promise of these materials and their nanotechnology applications (Bondarenko et al., 2013; Juganson et al., 2015). As concluded by the EU NanoSafety Cluster, the real concern rather than fragmentary observations on some hazards of exposure to ENMs, is the lack of systematic studies on adverse effects or exposure to ENMs (Savolainen et al., 2013). Since experimental testing is significantly constrained by time, financial burden, and ethical considerations (such as the principles of the 3Rs of animal testing, i.e. replacement, reduction, and refinement), the use of computational tools as alternative or compensation is expected to provide an efficient and inexpensive way of meeting the data requirements for the purpose of managing ENM risks (Raies and Bajic, 2016). Computational toxicology is seen as a potential tool to reduce the tension caused by the lag of evaluating nanosafety in respect to the rapid development of nanotechnology and nano-related innovation. Computational toxicology is emerging as a tool with active development and great potential (Reisfeld and Mayeno, 2012), and is able to create predictive power in the field of toxicology with the aid of modern computing and information technology (Kavlock et al., 2008; U.S. EPA, 2003).

Computational tools combined with powerful data-mining technologies, have been proposed to model chemical properties of soluble chemicals (Chen et al., 2014; Pavan et al., 2006; Tunkel et al., 2000), biological activities (Raies and Bajic, 2016), and species sensitivity distributions (SSDs) (Posthuma et al., 2002). The successful application of computational toxicology for soluble chemicals has promoted the expansion of these *in silico* approaches into the field of hazard identification of ENMs. Reliable computational tools can contribute to the supplementation of data for the gathering and evaluation of information as the first step of ENM hazard assessment recommended by the European Chemicals Agency (ECHA); or assist in the second step of hazard assessment (categorization and labeling of ENMs), by directly classifying ENMs into groups of different hazard (ECHA, 2011). For ENMs that meet the criteria of any of the hazard categories listed by ECHA, the use of the SSD method is helpful for deriving hazard threshold levels, e.g. predicted no effect concentration for the ecosystem as the last step of the ENM hazard assessment (ECHA, 2011). The information obtained on the basis of these steps is crucial for the qualitative risk characterization of ENMs. Thereupon, the structural characteristics that are identified by computational tools as governing toxicity may provide guidance for the safe-by-design of ENMs. However with these exciting promises in mind, challenges undoubtedly lie ahead as

this new research area is still in its infancy. This PhD research took the challenge and also the opportunity, aiming:

(i) To evaluate the currently existing literature data on metal-based ENMs for the use of computational toxicology in light of the safety assessment of ENMs;

(ii) To develop nano-(Q)SARs for the prediction and categorization of ENM hazard;

(iii) To derive SSDs and maximum acceptable environmental concentrations of metal-based ENMs as toxicity measures characterizing relevant risks.

To meet these research objectives, we have initially established an inventory of existing toxicity data of metal-based ENMs to selected organisms and identified data gaps as a preparation for ENM-related modeling (Chapter 2). The state-of-art of the (quantitative) structure–activity relationships for ENMs (nano-(Q)SARs) was reviewed regarding the availability of databases, the models developed up till now, the relevant descriptors commonly used, and on the basis of these advances, the options for interpretation of mechanisms of toxicity (Chapter 3). Later on, nano-SARs were developed for the categorization of ENM hazards to assist risk assessment and regulatory decision-making (Chapter 4). And finally, different SSDs were derived in Chapter 5 using currently available toxicity data for the generation of hazard threshold levels of ENMs.

With these approaches having been made, details related to each research objective of the thesis have been thoroughly discussed in the relevant chapters. To compare the developed models in this thesis with existing studies and also to provide implications for further advancing this new research frontier, some issues still need to be addressed based on the state-of-the-art of the application of computational toxicity in serving the hazard assessment of ENMs. The first issue standing out on this background is derived from the doubt of how well computational toxicology can heretofore assist ENM hazard assessment, from the prediction of ENM toxicity to the classification of ENM hazards, and to the derivation of hazard threshold levels as policy measures for the ENMs of concern. This issue subsequently leads to the further discussion of the situation of the constant struggle of data availability in ENM-related modeling, and to the key factors affecting nanotoxicity as indicated by the developed models for ENMs. Last but not least the challenges and outlook in this field are highlighted.

6.1 State-of-the-art of *in silico* models serving hazard assessment of ENMs

Given the limited availability and quality of existing data on nanotoxicity, doubt firstly arises about how well computational toxicology can contribute to the assessment of ENM hazards to date, including the discussion on the number and types of ENMs involved in the models; the potential applicability of these models in the assessment of ENM hazards; the descriptors used in the models; the information extracted for the safe-by-design of ENMs; the levels of the maximum acceptable concentrations of different ENMs; and the identified environmental risks of ENMs (if relevant information was presented in the underlying data sources). To answer these questions, a literature search of recent advances in the use of computational toxicology in developing *in silico* models for ENMs was performed. This was done by means of an Advanced Search in the Web of Science™ Core Collection on the 22th of February, 2017. The search was manually supplemented with relevant publications not included in the search records. The class of ENMs considered was restricted to metal and metal oxide ENMs. All relevant articles on the development of models for evaluating ENM hazards were selected and the reported models are reviewed and summarized in Tables 6.1 and 6.2.

6.1.1 Development of (Q)SARs and read-across models for metallic ENMs

As seen in Table 6.1, both regression and classification models predicting the biological activity profiles of metal-based ENMs have been developed. Ideally, a regression model is able to provide quantitative estimates for the hazardous effects of untested ENMs (or to untested species) and to fill in data gaps, which is fundamental for the evaluation of ENM toxicity. Classification models directly contribute to the categorization and labeling of ENMs. According to Table 6.1, 14 out of 22 of the studies originating from the literature review focused on the numerical prediction of ENM toxicity, and the rest of the studies presented classification models for the grouping of ENM hazards. Among these studies, a fair part of them aimed to predict the toxicity of metallic ENMs to *Escherichia coli* or to different cell lines; only three studies constructed models for other types of species (Chen et al., 2016; Kleandrova et al., 2014; Liu et al., 2013a). Categorical prediction of ENM toxicity could potentially serve the risk assessment of ENMs targeting a relatively broader spectrum of species given the current advances. For most of the *in silico* models, the datasets used are relatively small which probably poses major limitation on their potential applicability. Only two studies employed datasets of more than 100 ENMs (Chen et al., 2016; Kleandrova et al., 2014).

The frequently appearing descriptors in the models may encode important messages on ENM characteristics dominating relevant biological activities. This kind of messages

benefits both the hazard assessment and the safe-by-design of ENMs. Thus, the presented descriptors in existing models are summarized (see Table 6.1) and analyzed to discuss the role of different factors in influencing nanotoxicity. As for studies introducing multiple models or incorporating a big variety of descriptors, only main factors as highlighted by the authors are considered to avoid the impact of possible accidental correlations. The analysis show that some of the statistical models comprise merely theoretical descriptors; meanwhile the experimental parameters such as zeta potential, concentration of ENMs, aggregation parameter, size of the particles in media etc. are also found to be incorporated into other models. Subsequently, these descriptors are roughly labeled as belonging to one of three general types for further analysis: the intrinsic properties of the metal or metal oxide, the nano-specific characteristics of ENMs, and the dynamic changes of ENMs in media. The factors affecting ENM toxicity are further discussed here.

(i) Descriptors regarding the intrinsic properties of metal (oxide):

a. Surface catalytic properties and redox modifications related factors include: Wigner-Seitz radius, mass density, band gap energy, overlap of conduction band energy levels with the cellular redox potential, conduction band energy, average of the alpha and beta LUMO (lowest unoccupied molecular orbital) energies of the metal oxide, accessible surface area, absolute electronegativity of the metal and the metal oxide, aligned electronegativity, electronegativity, Mulliken's electronegativity of the cluster, S_2 (SiRMS-derived number of oxygen's atoms in a molecule, which was described by their electronegativity), S_3 (tri-atomic fragments[Me]–[O]–[Me] which were encoded by SiRMS-derived descriptors, encoding electronegativity), and metal electronegativity;

b. Characteristics related to the capability of ion and electron detachment and the activity of ions include: covalent index, cation polarizing power, atomization energy, metal oxide ionization energy, ionic index of metal cation, enthalpy of formation of metal oxide nanocluster representing a fragment of the surface, cationic charge, enthalpy of formation of a gaseous cation, charge of the metal cation corresponding to a given oxide, solubility, polarizability, molar refractivity, and polarization force;

(ii) The nano-specific descriptors employed in the developed models include:

a. The size of ENMs; and

b. Parameters characterizing the surface chemistry of ENMs, e.g., hydrophobicity of surface coating chemicals, surface-area-to-volume ratio, surface coating and charge, surface area, polar surface area;

Table 6.1. Summary of the state-of-the-art of developed (Q)SARs and read-across approaches for metal-based ENMs. Classification models are marked separately by means of an asterisk (*). N/A indicates that relevant information is not available

Reference	Indicated ENM characteristics in models	Theoretical descriptor	Experimental descriptor	ENMs	Tested organism	Data retrieved from
Liu et al., 2011*	Number of metal and oxygen atoms, molecular weight, atomization energy, group and period in the periodic table, size, isoelectric point, zeta potential, concentration	✓	✓	9 metal oxide ENMs	BEAS-2B cells	N/A
Zhang et al., 2012	Band gap energy, overlap of conduction band energy levels with the cellular redox potential (−4.12 to −4.84 eV), solubility	✓	✓	24 metal oxide ENMs	BEAS-2B cells, RAW 264.7 cells	N/A
Sizochenko et al., 2015	Mass density, molecular weight, aligned electronegativity, covalent index, cation polarizing power, Wigner-Seitz radius, surface area, surface-area-to-volume ratio, aggregation parameter, two-atomic descriptor of van-der-Waals interactions, tri-atomic descriptor of atomic charges, tetra-atomic descriptor of atomic charges, size in DMEM	✓	✓			Zhang et al., 2012
Liu et al., 2013b*	Atomization energy, atomic mass, size, conduction band energy, metal oxide ionization energy, electronegativity, ionic index of metal cation	✓				
Gajewicz et al., 2015a	Enthalpy of formation of metal oxide nanocluster representing a fragment of the surface, Mulliken's electronegativity of the cluster	✓		18 metal oxide ENMs	HaCaT cells	N/A
Luan et al., 2014*	Molar volume, polarizability, size	✓	✓	41 metallic ENMs	Mammalian cells	Multiple resources
Fourches et al., 2010*	Size, relaxivities, zeta potential	✓	✓	50 metallic ENMs	Endothelial cells, vascular smooth muscle cells, human HepG2 cells, RAW 264.7 cells	Shaw et al., 2008
Epa et al., 2012	Indicator variables of core material, surface coating, and surface charge	✓				

Table 6.1. (Continued)

Singh and Gupta, 2014	(i) Size, relaxivities, zeta potential; (ii) Oxygen percent, molar refractivity, polar surface area	✓	✓	(i) 44; (ii) 17 metallic ENMs	(i) Endothelial cells, vascular smooth muscle cells, human HepG2 cells, RAW 264.7 cells; (ii) *E coli*	Shaw et al., 2008; (ii) Puzyn et al., 2011
Sayes and Ivanov, 2010*	Size, concentration, size in phosphate buffered saline, size in water, zeta potential	✓	✓	24 TiO_2, 18 ZnO ENMs	Rat L2 lung epithelial cells, rat lung alveolar	N/A
Papa et al., 2015	Size, concentration, size in phosphate buffered saline, size in water	✓	✓		macrophages	Sayes and Ivanov, 2010
Pan et al., 2016	Molecular weight, cationic charge, mass percentage of metal elements, size, aggregation size	✓	✓	(i) 17; (ii) 18 metal oxide ENMs	(i) *E coli*; (ii) HaCaT cells	(i) Puzyn et al., 2011; (ii) Gajewicz et al., 2015
Gajewicz et al., 2015b*	Enthalpy of formation of a gaseous cation, Mulliken's electronegativity of the cluster	✓				
Sizochenko et al., 2014	(i) S_1, Wigner–Seitz radius, mass density, cation polarizing power, S_2, S_3, proportion of surface molecules to molecules in volume; (ii) S_1, Wigner–Seitz radius of oxide's molecule, mass density, covalent index of the metal ion, S_2, aggregation parameter	✓	✓			
Gajewicz et al., 2017	Enthalpy of formation of a gaseous cation, enthalpy of formation of metal oxide nanocluster representing a fragment of the surface, Mulliken's electronegativity of the cluster	✓				
Puzyn et al., 2011	Enthalpy of formation of a gaseous cation	✓		17 metal oxide ENMs	*E coli*	N/A
Mu et al., 2016	Polarization force, enthalpy of formation of a gaseous cation	✓				Puzyn et al., 2011
Kar et al., 2014	Charge of the metal cation corresponding to a given oxide, metal electronegativity	✓				
Pathakoti et al., 2014	Dark: absolute electronegativity of the metal and the metal oxide; Light: molar heat capacity, average of the alpha and beta LUMO (lowest unoccupied molecular orbital) energies of the metal oxide	✓				N/A

Table 6.1. (Continued)

Singh and Gupta, 2014	(i) Size, relaxivities, zeta potential; (ii) Oxygen percent, molar refractivity, polar surface area	✓	✓	(i) 44; (ii) 17 metallic ENMs	(i) Endothelial cells, vascular smooth muscle cells, human HepG2 cells, RAW 264.7 cells; (ii) *E coli*	Shaw et al., 2008; (ii) Puzyn et al., 2011
Chen et al., 2016* (Chapter 4)	Molecular polarizability, accessible surface area, solubility	✓		400; 450; 166 metallic ENMs	Various species	Chen et al., 2015; OCHEM
Kleandrova et al., 2014*	Molar volume, electronegativity, polarizability, size, hydrophobicity, polar surface area	✓	✓	229 metallic ENMs	Various species	Multiple resources
Liu et al., 2013a	Concentration, shell composition, surface functional groups, purity, core structure, and surface charge	✓	✓	82 ENMs including metal and metal oxide ENMs, dendrimer, polymeric etc.	Zebrafish embryo	NBI knowledgebase

The literature search was performed by means of an Advanced Search in the Web of Science™ Core Collection on the 22th of February, 2017. The query is ((((TS=(nano* AND metal)) AND (TS=(toxic*))) AND (TS=(quantitative *structure activity relationship) OR TS=(*QSAR) OR TS=(QNAR) OR TS=(predict*) OR TS=(computation*) OR TS=(model*)))), where the field tag TS refers to the topic of a publication; *E coli* - *Escherichia coli*; BEAS-2B - transformed bronchial epithelial cells; RAW 264.7 - murine myeloid cells; HaCaT - human keratinocyte cells; HepG2 cells – hepatocytes; S_1 - unbonded two-atomic fragments [Me]···[Me], which were encoded based on SiRMS-derived descriptors, describing the distance where potential reaches minimum at van der Waals interactions; S_2 - SiRMS-derived number of oxygen's atoms in a molecule, which was described by their electronegativity; S_3 - tri-atomic fragments[Me]–[O]–[Me], which were encoded by SiRMS-derived descriptors, encoding electronegativity; OCHEM- Online chemical modeling environment (Sushko et al., 2011); NBI Knowledgebase - Nanomaterial-Biological Interactions Knowledgebase.

(iii) The parameters indicating the dynamic changes of ENMs in media include:

a. Zeta potential;

b. Concentration of ENMs; and

c. Descriptors representing the dispersion and aggregation of ENMs in media, e.g., aggregation parameter, size in DMEM (Dulbecco's Modified Eagle's Medium), relaxivity (representing ENM magnetic properties), size in phosphate buffered saline, size in water, aggregation size.

Extraction of the general dependency of nanotoxicity on different factors may be of potential help for designing safe and environmentally benign ENMs. This kind of messages could be derived from the quantitative models for ENMs. Descriptors reported without explicit equations of predictive models cannot serve this purpose. As a result, despite the fact that various types of descriptors have been used in different *in silico* models, only a limited number of these parameters exhibited an explicit and unambiguous role in ENM-induced toxicity. The identified descriptors were roughly concluded here as concerning four aspects of the materials: the characteristics of ENMs *per se*, surface redox activity of metal oxides, ease of ion and electron detachment, and activity of the ion detached (see Figure 6.1). Some of the computational parameters may refer to multiply processes involved in the adverse effects triggered by metallic ENMs.

As can be seen from Figure 6.1, the hydrophobicity of ENM surface coatings and solubility of ENMs were shown to positively correlate with observed nanotoxicity. Other factors playing the same role in affecting nanotoxicity include the Wigner-Seitz radius and the electronegativity of metal oxides (χ_{oxide}) which reflect the surface redox activity of the metal or metal oxide; and the period in the periodic table of the ENM core metal, polarizability, and enthalpy of formation of metal oxide nanoclusters representing a fragment of the surface (ΔH_f^0), which indicates the ease of detachment of ions and electrons from ENMs. The Wigner-Seitz radius describes the available fraction of molecules on the surface of a nanocluster (Sizochenko et al., 2014). The χ_{oxide} characterizes the ability of atoms of metal oxides to attract electrons that contributes to the surface redox activities, and also relates to the leaching of ions from the surface of metal oxides (Gajewicz et al., 2015a). The period of the ENM metal represents information of atomic radii of the metal which is also associated with polarizability (Mahan and Subbaswamy, 1990).

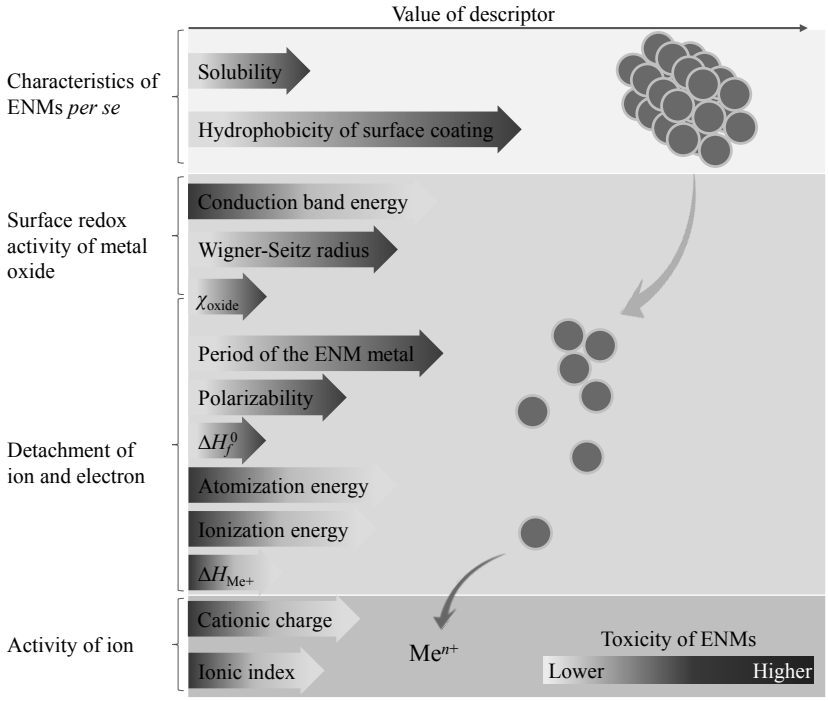


Figure 6.1. Generalization of the role of different factors in affecting the toxicity of metallic ENMs based on the state-of-the-art of nano-(Q)SARs and read-across models for ENMs. Me^{n+} represents the released ions from ENMs; ΔH_f^0 is the enthalpy of formation of metal oxide nanocluster representing a fragment of the surface; ΔH_{Me+} is the enthalpy of formation of a gaseous cation having the same oxidation state as that in the metal oxide structure; and χ_{cation} represents the electronegativity of the metal oxide.

On the other hand, the toxicity of ENMs tends to decrease with increased conduction band energy, atomization energy, ionization energy, ΔH_{Me+} (enthalpy of formation of a gaseous cation having the same oxidation state as the metal in the metal oxide structure), cationic charge, and ionic index. Zhang et al. (2012) have evidenced the strong correlation between the toxicity of Co_3O_4, Cr_2O_3, Ni_2O_3, Mn_2O_3, and CoO ENMs and the overlap of ENMs' conduction band energy with the cellular redox potential (-4.12 to -4.84 eV). The studied ENMs with conduction band energy out of the range failed to exhibit pro-oxidative and oxidative stress effects, with two exceptions ZnO and CuO ENMs. The exceptions could be explained by their relatively high solubility (Zhang et al., 2012). Decreasing atomization energy attributes to the decrease of the stability of metal oxides and corresponding increase

of reactivity (Liu et al., 2011). Ionization energy reflects the required amount of energy to remove the most loosely bound electron, a lower ionization energy thus indicates the easier detachment of electrons from the metal oxides (Bendary et al., 2013). ΔH_{Me+} describes the dissolution of ENMs without oxidation or reduction of ions, and the redox properties of metal oxides (Puzyn et al., 2011). Cationic charge was also found to be an important parameter in nano-QSARs (Pan et al., 2016). Cations (Me^{n+}) with smaller charges are considered more energetically favorable than cations of larger charges, which explains why the toxicity of metal oxides decreases in the order of $Me^{2+} > Me^{3+} > Me^{4+}$ (Puzyn et al., 2011). The ionic index of cations is associated with the affinity of metal ions for water molecules (measured by the hydration enthalpy); a lower hydration enthalpy means greater transport of metal ions across cellular membranes (Liu et al., 2013b). Notably, even though most of the employed descriptors characterize the intrinsic properties of the metal or metal oxides, several factors related to the characteristics of ENMs *per se* were also identified as affecting toxicity.

However, the role of some factors as concluded from developed models yielded conflicting results compared with experimental observations. For instance, the smooth muscle apoptosis (SMA) was modeled by means of the core material (I_{Fe3O4}), surface coating ($I_{dextran}$) and surface charge ($I_{surf.chg}$) of ENMs (Epa et al., 2012), and can be expressed as:

$$SMA = 2.26(\pm 0.72) - 10.73(\pm 1.05) \times I_{Fe_2O_3} - 5.57(\pm 0.98) \times I_{\text{dextran}} - 3.53(\pm 0.54) \times I_{\text{surf.chg}}$$

Therefore, based on this model it is obvious that a lower surface charge will result in higher apoptosis of smooth muscle cells. This, however, does not agree with some previous findings (Asati et al., 2010; EI Badawy et al., 2011; Schaeublin et al., 2011). Reportedly, the more negative citrate-Ag ENMs were the least toxic to gram-positive *bacillus*, whereas the positively charged Ag ENMs showed the strongest toxicity (EI Badawy et al., 2011). For Au ENMs, both the positively and negatively surface-charged Au ENMs were found to induce significant cellular mitochondrial stress other than the Au ENMs with neutral surface charge (Schaeublin et al., 2011). Another study of Asati et al. (2010) indicated that the surface charge of cerium oxide ENMs distinctly affects the internalization of ENMs by different cells, and the subsequent internal localization in cells which ultimately leads to the different toxicity profiles reported for cerium oxide ENMs. Meanwhile, the roles of some employed descriptors also conflict within or between independent studies. One example is the size of ENMs. The studies of both Luan et al. (2014) and Kleandrova et al. (2014) reported the diminution of ENM toxicity as a result of increasing ENM size. By contrast, based on the model developed by Liu et al. (2011), a larger size of ENMs was shown to lead to higher nanotoxicity. It was explained that indeed within the narrow domain of the dataset

(8–19 nm), toxicity may increase with increased primary size of ENMs. A linear model developed by Papa et al. (2015) also showed increased release of lactate dehydrogenase with the increment of the size of TiO_2 and ZnO ENMs (ranging from 20 to 70 nm). In addition, the particle size in phosphate buffered saline (PBS) and in water, indicating the aggregation behavior of ENMs in media, contributes oppositely to nanotoxicity as summarized from the models developed (Papa et al., 2015).

6.1.2 Development of SSDs for metal-based ENMs

The developed SSDs for metallic ENMs are summarized in Table 6.2. The state-of-the-art of the development of SSDs for metallic ENMs shows that Ag, Al_2O_3, Au, CeO_2, Cu, CuO, FeO_x, Silica, TiO_2, and ZnO ENMs have been commonly assessed for their adverse effects across different taxonomic groups. Compared to the diversity of ENMs involved in nano-(Q)SARs, the number of ENMs covered in SSD-related studies seems very limited. This may be because most of the derived SSDs grouped the materials solely based on their types (core material) without considering other structural characteristics. Thus, data of different ENMs with the same core was merged into the information of merely one type of ENMs. The exception is that, in the study of Garner et al. (2015) separate SSDs were presented for uncoated Ag and polyvinylpyrrolidone (PVP)-coated Ag ENMs. In Chapter 5 of this thesis, separate SSDs for metallic ENMs were developed considering different ENM characteristics, experimental conditions, and toxicity endpoints; separate SSDs were obtained for Ag ENMs grouped by surface coating, size, shape, and exposure duration; for CuO and ZnO ENMs grouped by size; for TiO_2 ENMs grouped by size and light exposure; and for Ag, CuO, TiO_2, and ZnO ENMs based on different toxicity endpoints (Chen et al., 2017). The limited variation in types of ENMs included in the SSDs is mostly due to the insufficient number of data of other type ENMs originated from experimental assays.

Nevertheless, the kinds of ENMs studied in the development of SSDs are indeed among the types that are largely found in the applications and products on the market. According to the study of Keller and Lazareva (2014), the 10 major ENMs (production of >100 t/year) used within the global economy are: Ag, Al_2O_3, CeO_2, Cu, Fe, SiO_2, TiO_2, and ZnO ENMs, carbon nanotubes, and nanoclays. An estimate of Bondarenko et al. (2013) on the annual production of ENMs showed an order with regard of production volume, from high to low, of SiO_2 (5500 t/year), TiO_2 (3000 t/year), ZnO (550 t/year) ENMs, carbon nanotubes (300 t/year), FeO_x (55 t/year), CeO_x (55 t/year), AlO_x (55 t/year), Ag ENMs (55 t/year), quantum dots (0.6 t/year), and fullerenes (0.6 t/year). Therefore, it seems like it is possible to perform safety evaluation of all the metallic ENMs that are produced in high amounts. Among these ENMs, Ag ENMs have relatively gained most research attention. Table 5.1 in the thesis showed that Ag ENMs have been tested on the highest number of species

considering the available data on LC50, EC50, LOEC, and NOEC. This enabled the development of SSDs for Ag ENMs separated by the different key factors described above. Further studies, ideally, should focus on other types of ENMs for the comprehensive evaluation of nanosafety. Meanwhile, besides the aquatic hazards of metallic ENMs, the potential risks brought by ENMs in other environmental compartments (e.g., air, soil) should also be considered. The implementation of these research needs however strongly depends on the quality of laboratory derived raw data. The increase of the quality of experimental data combined with robust uncertainty quantification will contribute to the improvement of the quality of SSDs.

The HC5s derived from the SSDs developed for different ENMs are compared as depicted in Figure 6.2. The HC5 values from Chapter 5 were taken from the SSDs of ungrouped Ag, CuO, TiO_2, and ZnO ENMs based on LC50 data and in case of CeO_2 ENMs based on EC50 data for comparison. As can be seen, Ag, TiO_2, and ZnO ENMs have relatively more estimates from the studies, which however also yielded much wider ranges of the reported HC5 values. The range of the HC5s of Ag ENMs shifted more to the left compared with that of the ZnO and TiO_2 ENMs, indicting the higher potential of toxic impacts of Ag ENMs on the environment. The HC5 values of silica and FeO_x ENMs are significantly higher than those of Ag ENMs. The median HC5 values of Au ENMs also indicated their mild toxicity compared with the toxicity of Ag ENMs. However, without the quantification of uncertainty it is hard to conclude whether the difference is significant.

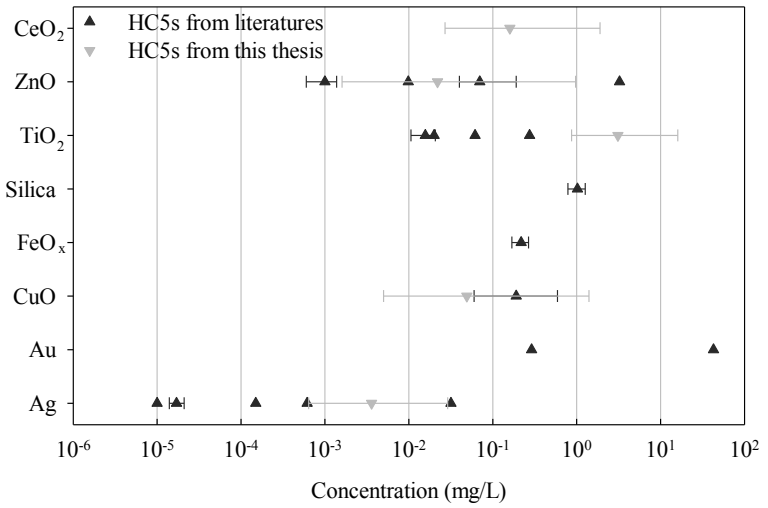


Figure 6.2. Estimated HC5s from SSDs (aquatic) for different types of ENMs. The relevant confidence intervals are also given (if available in the original publications).

Table 6.2. Summary of the state-of-the-art of the developed SSDs for metal and metal oxide ENMs. N/A indicates that relevant information is not available

Reference	Type of ENMs	Reported HC5s	Number of species in SSDs	Environmental compartment
Jacobs et al., 2016	TiO_2	N/A	31	Water
Wang et al., 2016a	FeO_x	0.218 (0.169-0.267) mg/L, 15-85% percentiles	12	Water
Kwak et al., 2016	Ag	0.03173 mg/L (acute toxicity); 0.000614 mg/L (chronic toxicity)	8 (acute toxicity); 5 (chronic toxicity)	Water
Coll et al., 2016	(i) Ag; (ii) TiO_2; (iii) ZnO	(i) 0.000017 (0.000014–0.000021) mg/L in freshwater, 8.2 (4.3–12.5) mg/kg in soil; (ii) 0.0157 (0.0106–0.0207) mg/L in fresh water, 91.1 (47.6–134.9) mg/kg in soil; (iii) 0.001 (0.0006–0.00138) mg/L in freshwater, 1.1 (0.6–1.6) mg/kg in soil, 95% confidence intervals	(i) 33 (water), 4(soil); (ii) 31 (water), 2 (soil); (iii) 21 (water), 7 (soil)	Water, soil
Wang et al., 2016b	Silica	1.023 (0.787-1.265) mg/L, 15-85% percentiles	8	Water
Mahapatra et al., 2015	Au	N/A	8 (water)	Water, soil
Semenzin et al., 2015	TiO_2	0.02 mg/L	34	Water
Adam et al., 2015	(i) ZnO; (ii) CuO	(i) 0.07 (0.04-0.19) mg/L; (ii) 0.19 (0.06-0.59) mg/L, 90% confidence intervals	(i) 12; (ii) 13	Water
Garner et al., 2015	(i) Ag; (ii) Cu; (iii) CuO; (iv) ZnO; (v) Al_2O_3; (vi) CeO_2; (vii) TiO_2	N/A	(i) Uncoated-Ag: 8, PVP-Ag: 6; (ii) 4; (iii) 5; (iv) 7; (v) 9; (vi) 7; (vii) 8	Water
Nam et al., 2015	Au	0.29 mg/L	7	Water
Botha et al., 2015	Au	42.78 mg/L	4	Water
Haulik et al., 2015	(i) Ag; (ii) TiO_2; (iii) ZnO	(i) 0.00015; (ii) 0.275; (iii) 3.246 mg/L	(i) 14; (ii) 11; (iii) 10	Water
Gottschalk et al., 2013	(i) Ag; (ii) TiO_2; (iii) ZnO	(i) 0.00001; (ii) 0.06151; (iii) 0.00985 mg/L	(i) 12; (ii) 18; (iii) 17	Water
Chen et al., 2017 (Chapter 5)	(i) Ag; (ii) CuO; (iii) ZnO; (iv) CeO_2; (v) TiO_2	HC5s were calculated for various SSDs (detailed information see Chapter 5)	Different hierarchies of species were used (detailed information see Chapter 5)	Water

The literature search was performed by means of an Advanced Search in the Web of Science™ Core Collection on the 22th of February, 2017. The query is (TS=(nano* AND *SSDs) OR TS=(nano* AND species sensitivity distributions)), where the field tag TS refers to the topic of a publication.

Additionally, a few studies have also presented the risk qualifications for metal-based ENMs along with the development of relevant SSDs, including Ag, Au, FeO_x, silica, TiO_2, and ZnO ENMs. For Ag ENMs, despite the estimated risks in surface water being shown by Haulik et al. (2015) to be below 0.001 (predicted environmental concentration divided by the HC5), the studies of both Gottschalk et al. (2013) and Coll et al. (2016) have reported significantly higher risk probabilities of respectively 0.7 and 0.038, which necessitates these materials to be studied in more depth with the highest priority. Risk coefficients of Ag ENMs in soil are calculated to be always <0.01 (Coll et al., 2016; Gottschalk et al., 2013). The risk coefficient of Ag ENMs in sewage treatment effluent is however as high as 39.7 (Gottschalk et al., 2013). Risk characterizations of TiO_2 ENMs in surface water and soil show that risks are relatively low in all studies except for the estimates reported by Coll et al. (2016) as being 0.03 and 0.013, respectively; the risk coefficient of TiO_2 ENMs in sewage treatment effluent is also relatively high (18.7). A marginal risk of ZnO ENMs in surface water (0.09) was indicted (Coll et al., 2016), whereas the risk coefficient of ZnO ENMs is again substantially higher (1.1) with respect to sewage treatment effluents (Gottschalk et al., 2013). For Au, FeO_x, and silica ENMs the derived risk probabilities are very low (Mahapatra et al., 2015; Wang et al., 2016a,b). In short, marginal risks are reported for Ag, TiO_2, and ZnO ENMs in surface water, and for TiO_2 ENMs in soil, while high environmental risks were identified for Ag, TiO_2, and ZnO ENMs in sewage treatment effluent.

6.2 The struggle of data availability

As was concluded from the state-of-the-art of the development of *in silico* models for metallic ENMs, an issue of vital importance in this field is the availability of reliable toxicity data. As described in Chapter 2 in the thesis, we have established a database assembling available and accessible data on the toxicity of metallic ENMs to algae, yeast, bacteria, protozoa, nematodes, crustacean, and fish. An analysis of the developed database (Figure 2.3) showed that most of the research attention was paid to merely a few species (e.g. *Pseudokirchneriella subcapitata*, *Staphylococcus aureus*, *Escherichia coli*, *Daphnia magna*, *Danio rerio*) and a few ENMs (e.g. Ag, CuO, TiO_2, ZnO ENMs). Despite the fact that in later Chapters (4 and 5) more data could be combined into the datasets to build nano-SARs and SSDs for ENMs, most of the information was still found available for only this limited number of species and ENMs. As shown in Chapter 4, sufficient data could be collected to build species-specific nano-SARs only for *Danio rerio*, *Daphnia magna*, *Pseudokirchneriella subcapitata*, *Escherichia coli*, and *Staphylococcus aureus*. In Chapter 5, SSDs could only be developed for Ag ENMs considering different factors and for CuO, TiO_2, and ZnO ENMs in some cases.

As a matter of fact, a total of 866 records of toxicity endpoints were collected in Chapter 2; in total 1061 toxicity records were made available within Chapter 4; and in Chapter 5 a total

of >1800 toxicity records could be retrieved from more than 300 publications or open access scientific reports. What has been noticed is that despite the continuing increase of the amount of data becoming available, the ENM-related modeling is still significantly constrained by the availability of experimental data (Gajewicz et al., 2017). Even with 1061 retrieved toxicity data in Chapter 3, nano-SARs could still only be developed by using descriptors characterizing the core information of ENMs, due to insufficient information on other ENM characteristics of importance (e.g. surface coating, shape, surface area, crystallinity). In Chapter 5, SSDs could be built solely for Ag ENMs roughly separated by the characteristics surface coating (uncoated, sodium citrate, PVP), size, and shape (spherical and non-spherical), given the nature of the >1800 toxicity records obtained.

For other nano-modeling studies as summarized in Table 6.1, in spite of the constantly increasing number of scientific resources from diverse nanosafety programs, only a relatively small number of datasets, such as those published by Puzyn et al. (2011) and Gajewicz et al. (2015a), were found to be repeatedly used in different modeling studies (Gajewicz et al., 2015a,b; 2017; Kar et al., 2014; Pan et al., 2016; Puzyn et al., 2011; Singh and Gupta, 2014; Sizochenko et al., 2014; Toropov et al., 2012). As shown in Table 6.2, the data points used for developing SSDs were also very limited. This leads to doubts about the suitability of existing nanotoxicity data in developing models for ENMs. As explained, data scarcity may result from data incompleteness and from inconsistency in reporting the characteristics of ENMs and relevant experimental information by independent studies. This in turn leads to the difficulty of comprehensively characterizing ENM structures for performing modeling and to the difficulty of separating ENMs according to different ENM characteristics or experimental conditions (Chen et al., 2016; 2017). In this context, availability of the vast majority of existing nanotoxicity data is greatly reduced and the use of this information in developing computational models for ENMs is severely prevented.

With limited available data on nanotoxicity, the developed models mostly incorporate descriptors representing only the ENM core, an approach that can also be used in the case of their corresponding bulk materials. As for further development of *in silico* models for ENMs, the ideal situation is to also involve comprehensive information on many of the other characteristics of ENMs such as surface chemistry, shape, dimensional aspects, crystallinity etc. for the better prediction and explanation of the biological activities of metallic ENMs (Chen et al., 2016). The use of parameters only characterizing ENM cores in models is by far not sufficient to address nano-specific toxicity in contrast with their bulk counterparts and to distinguish the structural differences of distinct ENMs with the same core. This requires a well-defined format for reporting the observed nanotoxicity, the experimental conditions, and the used ENMs. Thoroughly curated datasets of nanotoxicity are essential for modelers to carry out further researches. Therefore, here we propose that a

report of ENM toxicity for this specific purpose should properly describe at least the following information:

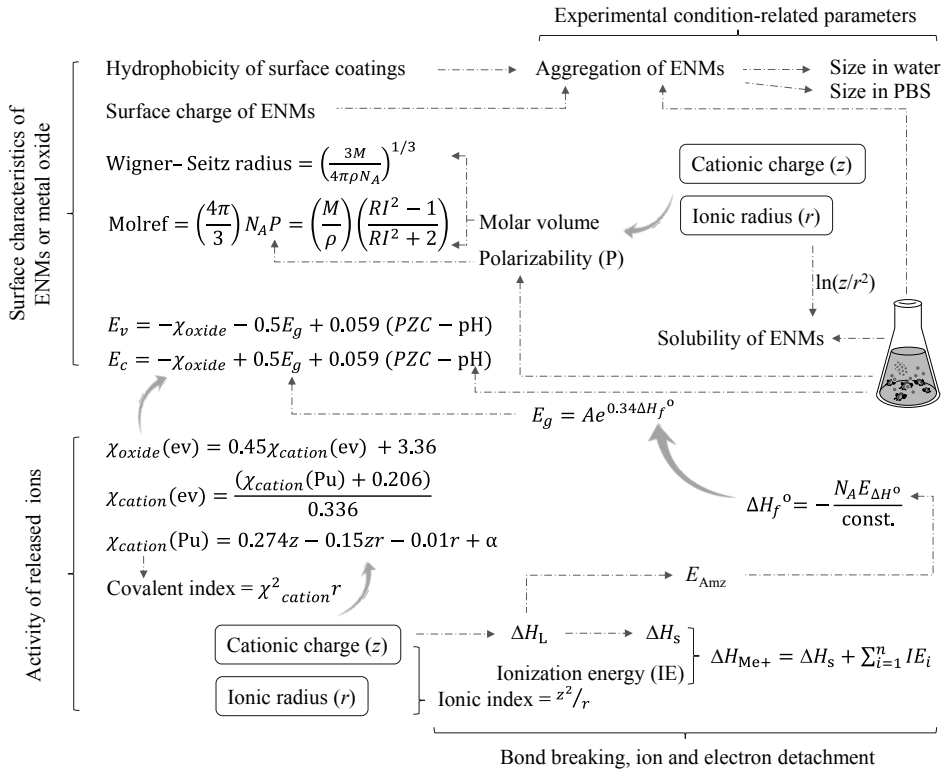


Figure 6.3. Profiling the toxicity of metal-based ENMs on the basis of identified descriptors. Dashed line indicates the simplified (mutual) correlation between the descriptors. The descriptors were roughly grouped as relating to the surface characteristics of ENMs or metal oxide, the activity of released ions, the bond breaking, ion and electron detachment, and the medium-related parameters. Molref - molar refractivity; M - molecular weight; ρ - density; N_A - Avogadro's number; RI - refractive index; PZC - point of zero charge; E_v - valence band energy; E_c - conduction band energy; E_g - band gap; χ_{oxide} - electronegativity of metal oxide; χ_{cation} - electronegativity of cation; E_{Amz} - atomization energy; ΔH_L - lattice energy; ΔH_s - enthalpy of sublimation; ΔH_{Me+} - enthalpy of formation of a gaseous cation having the same oxidation state as that in the metal oxide structure; $\Delta H_f{}^o$ - enthalpy of formation of metal oxide nanocluster representing a fragment of the surface; $E_{\Delta H}{}^o$ - energy associated with a single metal-oxygen bond in the metal oxide; PBS - phosphate buffered saline.

(i) Details of the tested organisms, e.g. taxonomic categorization, name of species, exposure route, life-stage or bacterial strain (for bacteria);

(ii) Conditions of the performed experiments, e.g. test guideline used (if available) and possible modifications of the test guideline, preparation of test medium, composition of the exposure medium, media pH, light condition, and time-dependent medium stability;

(iii) Information on the specific toxicity endpoints, e.g. observed biological effects, type of endpoint, experimental value of toxicity endpoint, and unit in which the endpoint is expressed; and

(iv) Characteristics of the ENMs tested, e.g. type of ENMs, composition of core, distribution of particle size, surface coating, purity, crystallinity, surface area, surface charge, shape, agglomerate size and material zeta potential in media, stability in test medium.

6.3 Profiling nanotoxicity on the basis of *in silico* models

The development of *in silico* models enabled the identification of factors of importance (represented by different descriptors) in affecting the toxicity of metallic ENMs. The hydrophobicity of surface coatings and surface charge of ENMs were shown to play an important role in determining nanotoxicity. These descriptors characterize the surface chemistry of metallic ENMs and are seen as nano-specific descriptors. The experimental conditions related parameters were also found in the reported models, including the solubility of ENMs, aggregation of ENMs, and relevant aggregated ENM size in the media. The rest of the commonly identified descriptors by nano-(Q)SARS or read-across models were seen as representing the intrinsic properties of the metal oxides, and generally belong to three groups that address different aspects of the material triggering adverse effects: descriptors describing the surface redox and catalytic properties of metal oxides; descriptors indicating the process of breaking of chemical bonds, detachment of ion and electron; and descriptors revealing the activity of ions released from ENMs.

For the sake of conciseness, a simplified explanation of the correlations of these descriptors is depicted in Figure 6.3. The conduction and valence band energies of metal oxide can be derived from their electronegativity, energy gap, point of zero charge, and pH of the media; the electronegativity of a metal oxide is derived from the electronegativity of the corresponding cation, which can be determined by the cationic charge and ionic radius based on the equations described in Figure 6.3 (Zhang et al., 2012). The cationic charge and

ionic radius likewise relate to the properties of metal oxides such as ionization energy (Ahrens, 1952), ionic index and atomization energy (Liu et al., 2013b), lattice energy (Puzyn et al., 2011), enthalpy of sublimation (Liu et al., 2013b; Puzyn et al., 2011), and the enthalpy of formation of a gaseous cation having the same oxidation state as in the metal oxide structure (Puzyn et al., 2011). Additionally, the cationic charge and ionic radius also relate to the polarizability and molar volume of the metal oxide (Mahan and Subbaswamy, 1990), and subsequently to other properties which are associated with these descriptors such as molar refractivity (Lide, 1998) and Wigner-Seitz radius (Sizochenko et al., 2014). Burello (2015) also classified the solubility of metal oxide ENMs in water and acidic media using the cationic charge and ionic radius. Therefore, it seems like the metal oxide ENMs which are able to release ions with smaller charge and larger ionic radius could induce higher toxicity to biota. That is to say, in general, within the same group of the periodic table, the larger the period that a metal belongs to (thus bigger atomic radius) the higher is the toxicity for the metal oxide ENMs formed by that metal; and within the same period in the table, metals on the left (thus smaller cationic charge) tend to form ENMs with higher toxicity compared with metals on the right. Meanwhile, metal oxide ENMs with low-valent metals may induce higher toxicity compared with ENMs composed of the same metal but of higher-valence. This corresponds with the study reported by Mu et al. (2016) which predicted the toxicity of 51 metal oxide ENMs to *Escherichia coli* (presented in a periodic table).

As previously explained in Chapter 1, it is commonly indicated that the release of ions and generation of reactive oxygen species (ROS) are two of the main mechanisms of metallic ENMs triggering toxicity, besides the possible direct steric hindrance caused by the particles *per se* and the ENMs acting as carriers of toxic chemicals (described as the Trojan-horse mechanism). In fact, both the detachment of ions or electrons from an ENM surface could lead to the formation of ROS. For instance, according to the Haber-Weiss-Fenon cycle (Gajewicz et al., 2015a; Stohs and Bagchi, 1995), Cu^{2+} could act as a catalyst for the formation of hydroxyl radicals ($OH^{\bullet}$), which subsequently leads to the generation of superoxide anion radicals ($O_2^{\bullet -}$):

$$O_2^{\bullet -} + Cu^{2+} \rightarrow O_2 + Cu^{+}$$

$$Cu^{+} + H_2O_2 \rightarrow Cu^{2+} + OH^{-} + OH^{\bullet}$$

$$O_2 + e \rightarrow O_2^{\bullet -}$$

Meanwhile, the detachment of an electron from the surface of TiO_2 ENMs (which could be activated by solar radiation) is also able to initiate a series of reactions leading to the formation of $OH^{\bullet}$ and $O_2^{\bullet -}$ (Kar et al., 2014):

$$TiO_2 \overset{h\nu}{\Rightarrow} TiO_2^+ + \bar{e}$$

$$\bar{e} + O_2 \rightarrow O_2^{\bullet -}$$

$$O_2^{\bullet -} + 2H^+ + \bar{e} \rightarrow H_2O_2$$

$$O_2^{\bullet -} + H_2O_2 \rightarrow OH^{\bullet} + OH^- + O_2$$

$$H^+ + H_2O \Rightarrow OH^{\bullet} + H^+$$

The generation of these ROS will disturb the cellular balance between the levels of oxidized and reduced species, and consequently provoke oxidative stress in cells (Gajewicz et al., 2015a). Thus, the intrinsic properties of a metal oxide (e.g., cationic charge and ionic radius) which are of significant importance for the possibility of electron transfer, bond breaking, and release of ions, seem to play a pivotal role in affecting the toxicity of ENMs. This is why doubt has arisen about whether the toxicity of metallic ENMs is nano-specific or comparable with that of corresponding dissolvable materials (Beer et al., 2012; Visnapuu et al., 2013; Xiu et al., 2012). However, undoubtedly, the other above-identified factors such as ENM surface chemistry, solubility of ENMs, and the experimental conditions are certainly able to alter the biological activity of metallic ENMs, by directly modifying the toxicity of the materials or by changing the bioavailability of ENMs for different species or cells (Fourches et al., 2010). In the study of Zhang et al. (2012), the solubility of metal oxide ENMs is one of the discriminating factors for classifying the observed toxicity. Solubility successfully explained the high toxicity of CuO and ZnO ENMs as the conduction band energy of the two ENMs has no overlap with the cellular redox potential (-4.12 to -4.84 eV). Observations of the nanotoxicity affected by ENMs shape were thereupon reported for ZnO nanospheres, nanosticks, and cuboidal submicron particles (Hua et al., 2014). The needle-shaped ZnO NPs were proven to be more toxic to *Phaeodactylum tricornutum* than other morphologically different NPs with equal solubility and ion release (Peng et al., 2011). Therefore, it seems that whether the toxicity induced by metallic ENMs should be considered as nano-specific is case-dependent.

Recently, a categorization framework of ENMs called the decision-making framework for the grouping and testing of nanomaterials (DF4nanoGrouping) was proposed based on the intrinsic material properties, system-dependent properties, and *in vitro* and *in vivo* effects of ENMs (Arts et al., 2015). This framework assigns ENMs into four main groups (MG) and determines to what extent the ENMs needs to be further evaluated. Specially, ENMs in MG 1 (soluble ENMs) are suggested to be handled by the read-across of the properties of dissolved materials from the bulk counterparts; ENMs in MG 4 (active ENMs) are advised

to be carefully evaluated and merit in-depth investigations in light of the risk assessment. ENMs in MG4 are for instance CeO_2 ENM-211, CeO_2 ENM-212, TiO_2 ENM-105, SiO_2 ENM‑acrylate, and SiO_2 ENM‑phosphate (Arts et al., 2016). Thus, based on this grouping strategy, the requirement on structural information of ENMs can be waived for the materials of MG 1. This kind of data is on the other hand of crucial importance for the "active" ENMs (Main Group 4), for the purpose of calculating nano-specific descriptors in case of generating *in silico* models for ENMs and for the purpose of grouping ENMs based on different properties in case of developing SSDs to diminish variabilities and levels of uncertainties.

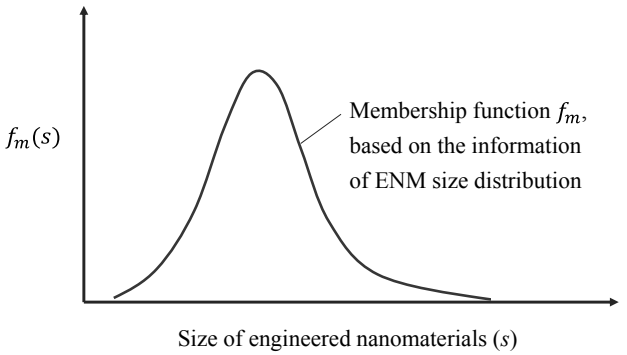


Figure 6.4. An explanation of considering the fuzzy set theory in handling the heterogeneity of ENM size for the computation of nano-specific descriptors.

6.4 Outlook

As previously addressed, one of the most fundamental issues in developing *in silico* models for ENMs is the availability and quality of laboratory derived data. For further experimental studies on nanotoxicity, providing comprehensive information according to standardized test protocols is of vital importance, together with widely accepted evaluation criteria for data quality. Meanwhile, maximizing the use of existing information seems realistic, practical, and favorable for this new frontier. One suggestion for this purpose is to transfer toxicity data between different endpoints with suitable assessment factors, which has been proven as a feasible way to obtain needed data given very limited available information. For example, in the study of Wang et al. (2016a,b), an assessment factor of 10 was used to transfer LC/EC25-50 to no observed effect concentrations; a factor of 2 for the LC/EC10-20; and a factor of 1 for other endpoints such as LOEC, LED, MIC, HONEC, and NOEC.

Likewise, this solution was also employed in different studies to overcome the problem of data scarcity (Coll et al., 2016; Gottschalk et al., 2013; Mahapatra et al., 2015). Even though uncertainty in doing so still remains debatable, this may be one of the most pragmatic ways of facing the current challenges of lack of toxicity data.

The structural complexity of ENMs has brought difficulty to computationally characterize the structure of ENMs in a comprehensive way. The incorporation of size information of ENMs into computational parameters also faces obstacles. An attempt to overcome this challenge is the study of Tämm et al. (2016) in which a set of novel, theoretical size-dependent nano-descriptors for ENMs was developed. However, the key problem is that the size of ENMs in reality is never a fixed value but rather a distribution of sizes. Preparing 100% homogeneous ENMs also does not seem possible in the near future. One proposed idea here is to adapt the calculation of nano-descriptors by combining them with fuzzy set theory. The fuzzy set theory permits the gradual assessment of the membership of elements in a set, instead of assigning an element into either one set or another (Zimmermann, 2010). Similarly, an ENM normally has a size distribution ranging, for example, 10-30 nm rather than a homogeneous size of 20 nm. Thus, if a descriptor (D_n) for a cluster of an ENM of size (s) can be expressed as:

$$D_n = f(s)$$

then the calculation of descriptors combined with fuzzy set theory (D_n') can be described as:

$$D_n' = \sum f_m(s) f(s)$$

when s is a discrete variable in $f(s)$, or

$$D_n' = \int_a^b f_m(s) f(s)\, d_s, a \le s \le b$$

when s is a continuous variable in $f(s)$; f_m is the membership function extracted from the information on the ENM size distribution (see Figure 6.4).

Another issue worth mentioning relates to the linking of structural characteristics of ENMs with their biological activities. As observed from Table 6.1, even though some of the studies constructed models solely on the basis of theoretical descriptors, the experimental descriptors such as zeta potential, concentration of ENMs, aggregation parameter, size in media etc. were also incorporated in other models. This agrees with the well-known fact

that the dynamic transformation of ENMs in media is able to alter the biological profiles of the materials. Thus in some cases toxicity information of ENMs can be poorly modeled without considering this transformation. However, dilemma situations arise as the safe-by-design approach of ENMs tends to favor the information of ENM safety purely based on their structures. For the next step, modeling and prediction of ENM behavior and transformation in different media (e.g. aggregation) could be considered based on ENM structural characteristics; and also the link of transformed characteristics of ENMs in the media to relevant biological activity. Different dose metrics in expressing the effective dose should be also taken into account for the modeling (Hua et al., 2016). Mass should not be the sole option in this context as nanotoxicity is influenced by many different physicochemical properties of ENMs (Oberdörster et al., 2007).

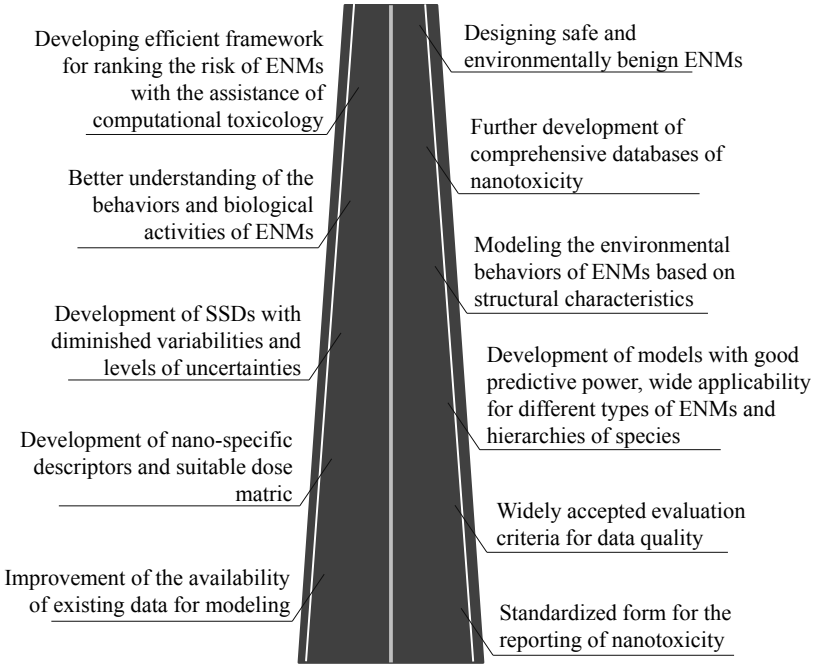


Figure 6.5. A roadmap indicating the future milestones of using computational toxicology in assisting the hazard assessment and safe-by-design of ENMs (drawn by G. Chen).

In the near future, the first milestone to be achieved regarding the use of computational toxicology in hazard assessment of ENMs should be a standardized form for reporting nanotoxicity (see Figure 6.5). Maximizing the use of existing data of nanotoxicity should

also be considered. Setting up widely accepted criteria is crucial for evaluating the quality of laboratory derived data for both existing and newly reported data. Development of novel nano-specific descriptors and incorporation of proper dose metrics are needed when performing modeling. The newly constructed nano-(Q)SARs and read-across models based on data with improved quality and availability are expected to have improved predictive power with broader applicability (suited for more types of ENMs and wider spectrum of species). The SSDs for deriving the maximum acceptable concentrations of ENMs are also expected to have diminished variabilities and levels of uncertainties. Meanwhile, linking the structural characteristics of ENMs to their environmental behavior and transformation is of great interest. Such work will provide further insight into the mechanisms underlying the biological profiles and environmental behavior of ENMs. In time, based on standardized criteria for reporting and evaluating nanotoxicity data, relevant databases with comprehensive information of all aspects will be developed. Upon these advances, construction of the framework ranking ENM hazard and associated risk aided by computational toxicology will highly contribute to the safe handling of ENMs and regulatory activities. Designing safe and environmentally benign ENMs supported by computational toxicology will also greatly benefit the minimization of risks brought by newly developed ENMs and the fast development of nanotechnology.

References

Adam N, Schmitt C, De Bruyn L, Knapen D, Blust R. Aquatic acute species sensitivity distributions of ZnO and CuO nanoparticles. Sci Total Environ. 2015, 526:233-42.

Ahrens LH. The use of ionization potentials Part 1. Ionic radii of the elements. Geochimica et Cosmochimica Acta. 1952, 2:155-169.

Arts JH, Hadi M, Irfan MA, Keene AM, Kreiling R, Lyon D, Maier M, Michel K, Petry T, Sauer UG, Warheit D, Wiench K, Wohlleben W, Landsiedel R. A decision-making framework for the grouping and testing of nanomaterials (DF4nanoGrouping). Regul Toxicol Pharmacol. 2015, 71:S1-27.

Arts JH, Irfan MA, Keene AM, Kreiling R, Lyon D, Maier M, Michel K, Neubauer N, Petry T, Sauer UG, Warheit D, Wiench K, Wohlleben W, Landsiedel R. Case studies putting the decision-making framework for the grouping and testing of nanomaterials (DF4nanoGrouping) into practice. Regul Toxicol Pharmacol. 2016, 76:234-61.

Asati A, Santra S, Kaittanis C, Perez JM. Surface-charge-dependent cell localization and cytotoxicity of cerium oxide nanoparticles. ACS Nano. 2010, 4:5321-31.

Beer C, Foldbjerg R, Hayashi Y, Sutherland DS, Autrup H. Toxicity of silver nanoparticles - nanoparticle or silver ion? Toxicol Lett. 2012, 208:286-92.

Bendary E, Francis RR, Ali HMG, Sarwat MI, El Hady S. Antioxidant and structure–activity relationships (SARs) of some phenolic and anilines compounds. Ann Agric Sci. 2013, 58:173-181.

Bondarenko O, Juganson K, Ivask A, Kasemets K, Mortimer M, Kahru A. Toxicity of Ag, CuO and ZnO nanoparticles to selected environmentally relevant test organisms and mammalian cells in vitro: a critical review. Arch Toxicol. 2013, 87:1181-200.

Botha TL, James TE, Wepener V. Comparative aquatic toxicity of gold nanoparticles and ionic gold using a species sensitivity distribution approach. J Nanomater. 2015, 2015:986902.

Burello E. Computational design of safer nanomaterials. Environ Sci: Nano. 2015, 2:454-462.

Chen G, Li X, Chen J, Zhang YN, Peijnenburg WJGM. Comparative study of biodegradability prediction of chemicals using decision trees, functional trees, and logistic regression. Environ Toxicol Chem. 2014, 33:2688-93.

Chen G, Peijnenburg WJGM, Kovalishyn V, Vijver MG. Development of nanostructure–activity relationships assisting the nanomaterial hazard categorization for risk assessment and regulatory decision-making. RSC Adv. 2016, 6:52227-52235.

Chen G, Peijnenburg WJGM, Xiao Y, Vijver MG. Developing species sensitivity distributions for metallic nanomaterials considering the characteristics of nanomaterials, experimental conditions, and different types of endpoints. Food Chem Toxicol. 2017, doi: 10.1016/j.fct.2017.04.003.

Chen G, Vijver MG, Peijnenburg WJGM. Summary and analysis of the currently existing literature data on metal-based nanoparticles published for selected aquatic organisms: Applicability for toxicity prediction by (Q)SARs. Altern Lab Anim. 2015, 43:221-40.

Coll C, Notter D, Gottschalk F, Sun T, Som C, Nowack B. Probabilistic environmental risk assessment of five nanomaterials (nano-TiO_2, nano-Ag, nano-ZnO, CNT, and fullerenes). Nanotoxicology. 2016, 10:436-44.

El Badawy AM, Silva RG, Morris B, Scheckel KG, Suidan MT, Tolaymat TM. Surface charge-dependent toxicity of silver nanoparticles. Environ Sci Technol. 2011, 45:283-7.

Epa VC, Burden FR, Tassa C, Weissleder R, Shaw S, Winkler DA. Modeling biological activities of nanoparticles. Nano Lett. 2012, 12:5808-12.

European Chemicals Agency (ECHA). Guidance on Information Requirements and Chemical Safety Assessment, Part B: Hazard Assessment, Version 2.1, European Chemicals Agency: Helsinki, Finland, 2011.

European Commission. Key Enabling Technologies. Available at https://ec.europa.eu/programmes/horizon2020/en/area/key-enabling-technologies. Accessed on 08.03.2017.

Fourches D, Pu D, Tassa C, Weissleder R, Shaw SY, Mumper RJ, Tropsha A. Quantitative nanostructure-activity relationship modeling. ACS Nano. 2010, 4:5703-12.

Gajewicz A, Cronin MT, 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. 2015b, 26:015701.

Gajewicz A, Jagiello K, Cronin MTD, 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. 2017, 4:346-358.

Gajewicz A, Schaeublin N, Rasulev B, Hussain S, Leszczynska D, Puzyn T, Leszczynski J. Towards understanding mechanisms governing cytotoxicity of metal oxides nanoparticles: hints from nano-QSAR studies. Nanotoxicology. 2015a, 9:313-25.

Garner KL, Suh S, Lenihan HS, Keller AA. Species Sensitivity Distributions for Engineered Nanomaterials. Environ Sci Technol. 2015, 49:5753-9.

Gottschalk F, Kost E, Nowack B. Engineered nanomaterials in water and soils: a risk quantification based on probabilistic exposure and effect modeling. Environ Toxicol Chem. 2013, 32:1278-87.

Haulik B, Balla S, Palfi O, Szekeres L, Jurikova T, Saly P, Bakonyi G. Comparative ecotoxicity of the nano Ag, TiO_2, and ZnO to aquatic species assemblages. Appl Ecol Env Res. 2015, 13:325-338.

Hua J, Vijver MG, Richardson MK, Ahmad F, Peijnenburg WJ. Particle-specific toxic effects of differently shaped zinc oxide nanoparticles to zebrafish embryos (Danio rerio). Environ Toxicol Chem. 2014, 33:2859-2868.

Hua J, Vijver MG, Chen G, Richardson MK, Peijnenburg WJ. Dose metrics assessment for differently shaped and sized metal-based nanoparticles. Environ Toxicol Chem. 2016, 35:2466-2473.

Jacobs R, Meesters JA, Ter Braak CJ, van de Meent D, van der Voet H. Combining exposure and effect modeling into an integrated probabilistic environmental risk assessment for nanoparticles. Environ Toxicol Chem. 2016, 35:2958-2967.

Juganson K, Ivask A, Blinova I, Mortimer M, Kahru A. NanoE-Tox: New and in-depth database concerning ecotoxicity of nanomaterials. Beilstein J Nanotechnol. 2015, 6:1788-804.

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. 2014, 107:162-9.

Kavlock RJ, Ankley G, Blancato J, Breen M, Conolly R, Dix D, Houck K, Hubal E, Judson R, Rabinowitz J, Richard A, Setzer RW, Shah I, Villeneuve D, Weber E. Computational toxicology—a state of the science mini review. Toxicol Sci. 2008, 103:14-27.

Keller AA, Lazareva A. Predicted Releases of Engineered Nanomaterials: From Global to Regional to Local. Environ Sci Technol Lett. 2014, 1:65-70.

Kleandrova VV, Luan F, González-Díaz H, Ruso JM, Speck-Planche A, Cordeiro MN. 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. 2014, 48:14686-94.

Kwak JI, Cui R, Nam SH, Kim SW, Chae Y, An YJ. Multispecies toxicity test for silver nanoparticles to derive hazardous concentration based on species sensitivity distribution for the protection of aquatic ecosystems. Nanotoxicology. 2016, 10:521-30.

Lide DR. CRC Handbook of Chemistry and Physics. CRC Press: Boca Raton, FL, 1998.

Liu R, Rallo R, George S, Ji Z, Nair S, Nel AE, Cohen Y. Classification NanoSAR development for cytotoxicity of metal oxide nanoparticles. Small. 2011, 7:1118-26.

Liu R, Zhang HY, Ji ZX, Rallo R, Xia T, Chang CH, Nel A, Cohen Y. Development of structure-activity relationship for metal oxide nanoparticles. Nanoscale. 2013b, 5:5644-53.

Liu X, Tang K, Harper S, Harper B, Steevens JA, Xu R. Predictive modeling of nanomaterial exposure effects in biological systems. Int J Nanomedicine. 2013a, 8:31-43.

Luan F, Kleandrova VV, González-Díaz H, Ruso JM, Melo A, Speck-Planche A, Cordeiro MN. Computer-aided nanotoxicology: assessing cytotoxicity of nanoparticles under diverse experimental conditions by using a novel QSTR-perturbation approach. Nanoscale. 2014, 6:10623-30.

Mahan GD, Subbaswamy KR. Local Density Theory of Polarizability. Plenum Press: New York, 1990.

Mahapatra I, Sun TY, Clark JR, Dobson PJ, Hungerbuehler K, Owen R, Nowack B, Lead J. Probabilistic modelling of prospective environmental concentrations of gold nanoparticles from medical applications as a basis for risk assessment. J Nanobiotechnology. 2015, 13:93.

Mu Y, Wu F, Zhao Q, Ji R, Qie Y, Zhou Y, Hu Y, Pang C, Hristozov D, Giesy JP, Xing B. Predicting toxic potencies of metal oxide nanoparticles by means of nano-QSARs. Nanotoxicology. 2016, 10:1207-14.

Nam SH, Shin YJ, Lee WM, Kim SW, Kwak JI, Yoon SJ, An YJ. Conducting a battery of bioassays for gold nanoparticles to derive guideline value for the protection of aquatic ecosystems. Nanotoxicology. 2015, 9:326-35.

NBI knowledgebase (Nanomaterial-Biological Interactions Knowledgebase). Available at http://nbi.oregonstate.edu/. Accessed on 08.02.2017.

Nel A, Xia T, Mädler L, Li N. Toxic potential of materials at the nanolevel. Science. 2006, 311:622-7.

Oberdörster G, Oberdörster E, Oberdörster J. Concepts of nanoparticle dose metric and response metric. Environ Health Perspect. 2007, 115:A290.

Pan Y, Li T, Cheng J, Telesca D, Zink JI, Jiang J. Nano-QSAR modeling for predicting the cytotoxicity of metal oxide nanoparticles using novel descriptors. RSC Adv. 2016, 6:25766-25775.

Papa E, Doucet JP, Doucet-Panaye A. Linear and non-linear modelling of the cytotoxicity of TiO_2 and ZnO nanoparticles by empirical descriptors. SAR QSAR Environ Res. 2015, 26:647-65.

Pathakoti K, Huang MJ, Watts JD, He X, Hwang HM. 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. 2014, 130:234-40.

Pavan M, Worth AP, Netzeva TI. Review of QSAR Models for Bioconcentration, European Commission, Joint Research Centre, Ispra, Italy, EUR 22327 EN, 2006.

Peng X, Palma S, Fisher NS, Wong SS. Effect of morphology of ZnO nanostructures on their toxicity to marine algae. Aquat Toxicol. 2011, 102:186-96.

Posthuma L, Traas TP, Suter GW. 2002. General introduction to species sensitivity distributions. In: Posthuma L, Suter GW, Traas TP. Eds., Species sensitivity distribution in ecotoxicology, Lewis, Boca Raton, FL, USA, pp. 3–10.

Puzyn T, Rasulev B, Gajewicz A, Hu X, Dasari TP, Michalkova A, Hwang HM, Toropov A, Leszczynska D, Leszczynski J. Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles. Nat Nanotechnol. 2011, 6:175-8.

Raies AB, Bajic VB. In silico toxicology: computational methods for the prediction of chemical toxicity. Wiley Interdiscip Rev Comput Mol Sci. 2016, 6:147-172.

Reisfeld B, Mayeno AN. What is computational toxicology? Reisfeld B, Mayeno AN, editors. Methods Mol Biol. 2012, 929, 3–7.

Savolainen K, Backman U, Brouwer D, Fadeel B, Fernandes T, Kuhlbusch T, Landsiedel R, Lynch I, Pylkkänen L. Nanosafety in Europe 2015-2025: Towards Safe and Sustainable Nanomaterials and Nanotechnology Innovations. Finnish Institute of Occupational Health: Helsinki, Finnland, 2013.

Sayes C, Ivanov I. Comparative Study of Predictive Computational Models for Nanoparticle-Induced Cytotoxicity. Risk Anal. 2010, 30:1723-34.

Schaeublin NM, Braydich-Stolle LK, Schrand AM, Miller JM, Hutchison J, Schlager JJ, Hussain SM. Surface charge of gold nanoparticles mediates mechanism of toxicity. Nanoscale. 2011, 3:410-20.

Semenzin E, Lanzellotto E, Hristozov D, Critto A, Zabeo A, Giubilato E, Marcomini A. Species sensitivity weighted distribution for ecological risk assessment of engineered nanomaterials: the n-TiO_2 case study. Environ Toxicol Chem. 2015, 34:2644-59.

Shaw SY, Westly EC, Pittet MJ, Subramanian A, Schreiber SL, Weissleder R. Perturbational profiling of nanomaterial biologic activity. Proc Natl Acad Sci U S A. 2008, 105:7387-92.

Singh KP, Gupta S. Nano-QSAR modeling for predicting biological activity of diverse nanomaterials. RSC Adv. 2014, 4:13215-13230.

Sizochenko N, Rasulev B, Gajewicz A, Kuz'min V, Puzyn T, Leszczynski J. From basic physics to mechanisms of toxicity: the "liquid drop" approach applied to develop predictive classification models for toxicity of metal oxide nanoparticles. Nanoscale. 2014, 6:13986-93.

Sizochenko N, Rasulev B, Gajewicz A, Mokshyna E, Kuz'min VE, Leszczynski J, Puzyn T. Causal inference methods to assist in mechanistic interpretation of classification nano-SAR models. RSC Adv. 2015, 5:77739-45.

Stohs SJ, Bagchi D. Oxidative mechanisms in the toxicity of metal ions. Free Radic Biol Med. 1995, 18:321-36.

Sushko I, Novotarskyi S, Körner R, Pandey AK, Rupp M, Teetz W, Brandmaier S, Abdelaziz A, Prokopenko VV, Tanchuk VY, Todeschini R, Varnek A, Marcou G, Ertl P, Potemkin V, Grishina M, Gasteiger J, Schwab C, Baskin II, Palyulin VA, Radchenko EV, Welsh WJ, Kholodovych V, Chekmarev D, Cherkasov A, Aires-de-Sousa J, Zhang QY, Bender A, Nigsch F, Patiny L, Williams A, Tkachenko V, Tetko IV. Online chemical modeling environment (OCHEM): web platform for data storage, model development and publishing of chemical information. J Comput Aided Mol Des. 2011, 25:533-54.

Tämm K, Sikk L, Burk J, Rallo R, Pokhrel S, Mädler L, Scott-Fordsmand JJ, Burk P, Tamm T. Parametrization of nanoparticles: development of full-particle nanodescriptors. Nanoscale. 2016, 8:16243-16250.

Toropov AA, Toropova AP, Benfenati E, Gini G, Puzyn T, Leszczynska D, Leszczynski J. Novel application of the CORAL software to model cytotoxicity of metal oxide nanoparticles to bacteria Escherichia coli. Chemosphere. 2012, 89:1098-102.

Tunkel J, Howard PH, Boethling RS, Stiteler W, Loonen H. Predicting ready biodegradability in the Japanese ministry of international trade and industry test. Environ Toxicol Chem. 2000, 19:2478-85.

U.S. EPA. A framework for a computational toxicology research program, 2003, Washington, D.C. EPA600/R-03/65.

Visnapuu M, Joost U, Juganson K, Künnis-Beres K, Kahru A, Kisand V, Ivask A. Dissolution of silver nanowires and nanospheres dictates their toxicity to Escherichia coli. Biomed Res Int. 2013, 2013:819252.

Wang Y, Deng L, Caballero-Guzman A, Nowack B. Are engineered nano iron oxide particles safe? an environmental risk assessment by probabilistic exposure, effects and risk modeling. Nanotoxicology. 2016a, 10:1545-1554.

Wang Y, Kalinina A, Sun T, Nowack B. Probabilistic modeling of the flows and environmental risks of nano-silica. Sci Total Environ. 2016b, 545-546:67-76.

Xiu ZM, Zhang QB, Puppala HL, Colvin VL, Alvarez PJ. Negligible particle-specific antibacterial activity of silver nanoparticles. Nano Lett. 2012, 12:4271-5.

Zimmermann H-J. Fuzzy set theory. Wiley Interdiscip Rev Comput Stat. 2010, 2:317-332.

Zhang H, Ji Z, Xia T, Meng H, Low-Kam C, Liu R, Pokhrel S, Lin S, Wang X, Liao YP, Wang M, Li L, Rallo R, Damoiseaux R, Telesca D, Mädler L, Cohen Y, Zink JI, Nel AE. Use of metal oxide nanoparticle band gap to develop a predictive paradigm for oxidative stress and acute pulmonary inflammation. ACS Nano. 2012, 6:4349-68.

Summary

Nanotechnology is seen as a revolutionary technology which greatly benefits the world economy. However, as usual there is a tension between the need to manufacture new nanomaterials with desired properties, and the need to protect the environment and human beings from the potential risks associated. The lag between the time needed to evaluate the safety of engineered nanomaterials (ENMs) and the rapid development of nanotechnology has already caused concerns about the safe use of ENMs. Assessing the risks of ENMs solely on the basis of experimental assays is time-consuming, resource intensive, and constrained by ethical considerations (such as the principles of the 3Rs of animal testing, i.e. replacement, reduction, and refinement). The adoption of computational toxicology in this field is a high priority. Computational toxicology is able to contribute to the prediction of the extent of toxic effects of untested ENMs, to the hazard categorization and labeling of ENMs, and to the establishment of hazard threshold values that are sufficiently protecting the ecosystem with respect to the ENMs of concern. These three steps are listed by the European Chemicals Agency (ECHA) as the three elements in evaluating the hazards of ENMs. A comprehensive hazard assessment for ENMs is essential for both the risk characterization and the safe-by-design of nanomaterials.

To facilitate the use of computational toxicology in assisting the hazard assessment of ENMs, the research of this thesis started from the integration and evaluation of existing available and accessible data regarding the toxicity of metal-based ENMs to selected organisms (**Chapter 2**). A database of 886 records was developed, containing information on bacteria, algae, yeast, protozoa, nematode, crustacean, and fish; and on ENMs composed of metals, metal oxides, nanocomposites, and quantum dots. The analysis indicated that Ag ENMs are the most widely studied ENMs, together with TiO_2 and ZnO ENMs. *Daphnia magna*, *Escherichia coli*, and *Pseudokirchneriella subcapitata* are the most frequently tested species in the database. Biological effects investigated for each group of organism were analyzed, and the types of ENMs and species in the database were described in as much detail as possible. ENMs were classified into different hazard categories adhering to the EU Directive 93/67/EEC.

Following up the data integration and evaluation, the state-of-the-art of the development of (quantitative) structure–activity relationships for ENMs (nano-(Q)SARs) was reviewed in **Chapter 3**. Issues concerning the sources of data for modeling, existing nano-(Q)SARs, and mechanistic interpretation were discussed

and an outlook on the further development of this field was presented. The analysis showed that cellular uptake of ENMs by different cells and the toxicity to *Escherichia coli* are the main focus of nano-(Q)SAR modeling. Models were developed for both quantitative and categorical predictions of the biological activities of ENMs based on different data mining approaches. As could be concluded from the identified descriptors, lipophilicity and hydrogen bonding capacity of surface modifiers were found to be of most significant importance for the cellular uptake of ENMs. The released ions and generation of oxidative stress are seen as driving factors in causing nanotoxicity in some cases; nano-specific properties such as surface chemistry, size are also believed to play a role. Similar to chapter 2, also here we saw the problem of data scarcity and data quality. The characterization of ENM structures and the consideration of dynamic transformations of ENMs in the exposure medium in modeling should also be carefully handled.

Based on the identified research gaps on nano-(Q)SARs, in **Chapter 4** the nano-SARs for the categorization of ENM hazards were built on the basis of the retrieval of existing toxicity data. The global nano-SARs across species in case study I (LC50 data, 320 ENMs in training set and 80 ENMs in test set) and III (MIC data, 133 ENMs in training set and 33 ENMs in test set) yielded reasonable accuracies (above 70%). Species-specific nano-SARs were also constructed for *Danio rerio*, *Daphnia magna*, *Pseudokirchneriella subcapitata*, and *Staphylococcus aureus* with high predictability. The molecular polarizability, accessible surface area, and solubility of ENMs were identified in the models that were built as predominantly influencing the toxicity of metallic ENMs. The study contributes to the classification and labeling of metallic ENMs for regulatory purposes.

Once an ENM is classified in one of the hazard classes or categories listed by ECHA, a risk characterization for the ENM is required. This necessities the derivation of threshold levels for ENMs in order to compare with relevant exposure levels and to quantify associated risks. In case of generic risk assessment, the 5th percentile (HC5) of the species sensitivity distributions (SSDs) is commonly used for this comparison. **Chapter 5** therefore focused on the development of SSDs for metallic ENMs with the explicit consideration of the characteristics of ENMs, experimental conditions, and different types of endpoints. Based upon a sufficient number of data entries, separate SSDs could only be built for Ag ENMs based on the characteristics surface coating, size, shape, and exposure duration. Separate SSDs were also developed to determine whether and to what extent the shape of the SSD curve alters and the resulting

HC5s varies based on different toxicity endpoints. As could be concluded from the developed SSDs, the PVP- and sodium citrate coatings were found to enhance the toxicity of Ag ENMs; for Ag ENMs with different size ranges, differences in behavior and/or effects were only observed at high exposure concentrations; the SSDs of Ag ENMs separated by either shape or exposure duration were all nearly identical. Meanwhile, crustaceans were found to be the most vulnerable group to metallic ENMs.

In conclusion, our study has expanded the use of computational toxicology in hazard assessment with regard to the safe handling of ENMs. The results obtained contribute to the integration and evaluation of toxicity data, the identification of research gaps on ENM-related modeling, and the development of nano-SARs and SSDs for metallic ENMs. Despite the uncertainties that are associated with our results, as mainly due to limited data quality and availability, we managed to take this field one step forwards and contribute to better-informed regulatory decisions of ENMs. To enable the next step to be made, it is essential that research in the relevant fields more strictly adhere to the guidance that has been issued regarding proper reporting of scientific data on the fate and effects of ENMs. This will allow for efficient data curation and proper comparison of experimental data.

Samenvatting

Nanotechnologie wordt gezien als een revolutionaire technologie waarvan de wereldeconomie zal profiteren. Echter, zoals gebruikelijk bij de introductie van nieuwe technologieën, is er een spanningsveld tussen de noodzaak om nieuwe nanomaterialen te produceren met gewenste eigenschappen en de noodzaak om het milieu en de mens te beschermen gelet op de potentiële risico's die met de technologie samenhangen. Er is een grote tijds-uitdaging tussen de snelle ontwikkeling van de nanotechnologie en de langere tijd die nodig is om de veiligheid van ontwikkelde nanomaterialen (ENMs) te evalueren. Het beoordelen van de risico's van ENMs uitsluitend op basis van experimentele testen is tijdrovend, duur, en beperkt door ethische overwegingen (zoals het principe van de 3Rs van dierproeven, d.w.z. vervanging, reductie en verfijning). Het implementeren van de computationele toxicologie voor ENMs heeft dan ook hoge prioriteit. Computationele toxicologie kan bijdragen aan de voorspelling van de mate van toxische effecten van niet-geteste ENMs, de categorisatie van risico's van ENMs en de vaststelling van drempelwaarden die het ecosysteem voldoende beschermen tegen ENMs. Deze drie stappen worden vermeld door het European Chemicals Agency (ECHA) als de drie elementen bij het evalueren van de risico's van ENMs. Een uitgebreide risicobeoordeling voor ENMs is essentieel voor zowel de risico-karakterisering alsook voor het safe-by-design ontwerpen van producten waarin ENMs worden gebruikt.

Om de computationele toxicologie te gebruiken bij het ondersteunen van de risicobeoordeling van ENMs, is het onderzoek beschreven in dit proefschrift gestart met de integratie en evaluatie van de beschikbare en toegankelijke gegevens over de toxiciteit van metaalhoudende ENMs (hoofdstuk 2). Een database van 886 records is ontwikkeld met informatie over de toxiciteit van ENMs voor bacteriën, algen, gisten, protozoa, nematodes, schaaldieren en vis; alsmede van alle metaalhoudende ENMs die bestaan uit metalen, metaaloxiden, nano-komposieten en kwantumdots. Uit de analyse blijkt dat Ag-gebaseerde ENMs de meest bestudeerde ENMs zijn, samen met ENMs die bestaan uit TiO_2 en ZnO. *Daphnia magna*, *Escherichia coli* en *Pseudokirchneriella subcapitata* zijn de meest geteste soorten in de database. De ENMs in de database werden ingedeeld in verschillende risico-categorieën volgens de EU-richtlijn 93/67/EEC.

Na de data-integratie en -evaluatie is de state-of-the-art van de ontwikkeling van (kwantitatieve) structuur-activiteitrelaties voor ENMs (nano-(Q)SARs) onderzocht in hoofdstuk 3. Problemen over de data en gegevensbronnen,

bestaande nano-(Q)SARs en mechanistische interpretatie werden besproken en een vooruitblik op de verdere ontwikkeling van dit veld werd gepresenteerd. De analyse toonde aan dat cellulaire opname van ENMs in verschillende celculturen en de toxiciteit voor *Escherichia coli* de belangrijkste focus zijn binnen de nano-(Q)SAR-modellering. Modellen werden ontwikkeld voor zowel kwantitatieve als categorische voorspellingen van de biologische activiteiten van ENMs gebruikmakend van verschillende data 'mining' benaderingen. Zoals uit de geïdentificeerde model beschrijvingen kon worden geconcludeerd, bleken lipofiele eigenschappen en waterstofbindingscapaciteit van functionele groepen aan het oppervlak van de ENMs van het grootste belang voor de cellulaire opname van ENMs. De vrijgekomen ionen en de generatie van oxidatieve stress worden in sommige gevallen als belangrijkste factoren beschouwd bij het veroorzaken van nanotoxiciteit. Verder wordt aangetoond dat nano-specifieke eigenschappen zoals oppervlaktechemie en de grootte van de deeltjes een belangrijke rol spelen bij de toxiciteit van ENMs. Evenals bij hoofdstuk 2, zagen we hier eveneens het probleem van de data schaarste en datakwaliteit. Tenslotte wordt geconcludeerd dat de gedetailleerde karakterisering van ENM-structuren en het karakteriseren van transformaties van ENMs in het blootstellingsmedium ook nauwgezet dienen te worden meegenomen in de modellering.

Gegeven de eerder geconstateerde beperkingen met betrekking tot de ontwikkeling van nano-(Q)SARs, zijn in hoofdstuk 4 nano-SARs ontwikkeld voor de categorisering van ENM-risico's op basis van bestaande toxiciteitsgegevens. De robuuste nano-SARs in case study I (LC50 data, 320 ENMs in de trainings-set en 80 ENMs in de test-set) en in case study III (MIC data voor de invloed van 133 ENMs op de remming van de activiteit van bacteriën in de trainings-set en 33 ENMs in de test-set) leverden voorspellingen op met nauwkeurigheden tot 70%. Daarnaast zijn soort-specifieke nano-SAR's ontwikkeld voor *Danio rerio*, *Daphnia magna*, *Pseudokirchneriella subcapitata* en *Staphylococcus aureus* met hoge accuratesse. De moleculaire polarisatie, oppervlakte grootte en oplosbaarheid van ENMs werden geïdentificeerd als zijnde de parameters die de toxiciteit van metaalhoudende ENMs het sterkste beïnvloeden. Deze studie draagt bij aan de classificatie en etikettering van metaalhoudende ENMs voor regelgevende doeleinden.

Zodra een ENM is ingedeeld in één van de door ECHA vermelde risico-klassen of categorieën, is een verdere risico-karakterisering vereist. Dit vereist de afleiding van drempelwaarden voor ENMs die dan kunnen worden vergeleken met relevante blootstellingsniveaus, om uiteindelijk de bijbehorende risico's te kwantificeren. Bij een generieke risicobeoordeling wordt het 5^{e} percentiel (HC5)

van de soortgevoeligheidsverdeling (SSDs) vaak gebruikt. Hoofdstuk 5 concentreerde zich derhalve op de ontwikkeling van SSDs voor metaalhoudende ENMs, met expliciete overweging van de kenmerken van ENMs, experimentele omstandigheden, en verschillende soorten met hun toxiciteits-eindpunten. Alleen voor Ag-houdende ENMs konden SSDs worden ontwikkeld op basis van de eigenschappen van de coating, de grootte en de vorm van de deeltjes, en de blootstellingsduur. Er werden ook afzonderlijke SSDs ontwikkeld om te bepalen *of* en in *welke mate* de vorm van de SSD-curve verandert en de resulterende HC5's variëren op basis van verschillende toxiciteits-eindpunten. Zoals uit de ontwikkelde SSDs kon worden geconcludeerd, bleken de PVP- en natriumcitraat-coatings de toxiciteit van Ag-houdende ENMs te verhogen; voor Ag-houdende ENMs met verschillende afmetingen, werden verschillen in gedrag in het blootstelingsmedium en/of effecten alleen waargenomen bij hoge blootstellingsconcentraties; de vorm van de Ag-houdende ENMs en de blootstellingsduur van de testen deed de SSD niet veranderen. Kreeftachtigen bleken de meest gevoelige groep van organismen te zijn voor de metaalhoudende ENMs.

Concluderend kan gesteld worden dat het onderzoek dat in dit proefschrift wordt beschreven, bijdraagt aan de uitbreiding van de toepassing van de computationele toxicologie in risicobeoordeling, en wel specifiek voor het inschatten van de milieurisico's van ENMs. De verkregen resultaten dragen bij aan de integratie en evaluatie van toxiciteitsgegevens, de identificatie van onderzoeksprioriteiten bij ENM-gerelateerde modellering, en aan de ontwikkeling van nano-(Q)SARs en SSDs voor metaalhoudende ENMs. Ondanks de onzekerheden die samenhangen met onze resultaten, veroorzaakt door de beperkte data kwaliteit en beschikbaarheid, slaagden we erin om dit onderzoeksveld een stap voorwaarts te brengen en dragen we bij aan de verbetering van regelgevende beslissingen voor ENMs. Om een significante vervolgstap te kunnen maken, is het essentieel dat de onderzoekers in het veld van de nanotoxicologie zich strikt houden aan de richtlijnen die zijn opgesteld voor de accurate rapportage en onderbouwing van wetenschappelijke gegevens over het gedrag en effecten van ENMs. Dit zorgt voor efficiënte data-curatie en voor de mogelijkheid om experimentele data onderling te kunnen vergelijken.

论文概要

纳米科技是近年来新兴的学科领域，被视为推动全球经济发展的新引擎。但是在不断研发与制造特定性能纳米材料的同时，使用这些材料所带来的潜在风险同样不可忽视。迄今为止，相较于纳米科技的迅猛发展，对纳米材料安全性的评估研究却仍然滞后，仍然需要引起警惕。对已有及新纳米材料安全性进行全面、综合的评估需要充足的数据信息，如果只依靠传统的方式，通过试验来获取信息很耗时耗力，也受限于动物试验的伦理问题（如以减少、替代以及优化为核心的 3R 动物实验原则）。因此近年来，很多研究者将注意力转移到应用计算毒理学技术辅助、支持纳米材料的安全性评估研究上。计算毒理学可应用于评估纳米材料的毒性作用，归类、标记纳米材料的危险性，以及估算纳米材料的生态风险阈值。此三部分内容被欧洲化学品管理局列为评价纳米材料危险性的三要素。对已有纳米材料的危险性进行综合的、全方位的评估既有助于进行相关的环境和人体健康风险评定，又有助于研发更安全的新型纳米材料。

为应用计算毒理学方法评估、分析纳米材料的安全性，本论文着手于收集、整理金属纳米材料的生态毒理数据（**第二章**）。此部分研究从现有文献中收集整理了 886 条纳米材料的生态毒理试验数据。数据集中所涉及的受试生物包括细菌、藻类、酵母、原生动物、线虫类、甲壳类以及鱼类，所涉及的纳米材料包括金属类和金属氧化物类纳米材料、纳米复合材料以及量子点。对数据集的分析表明，相较于其它纳米材料，有关银纳米材料的生态毒理数据最多，纳米氧化钛和氧化锌次之。在所涉及的受试生物中，有关大型溞、大肠杆菌和羊角月牙藻的研究最为广泛。此章研究分析整理了试验中纳米材料对各种生物的毒性效应。所整理的数据集涵盖了可收集到的有关纳米材料和受试生物的各种信息。参照欧盟委员会划分的污染物对水生生物的风险等级标准(EU Directive 93/67/EEC)，研究还对数据集中纳米材料的危险性进行了分类。

在数据收集整理之后，本论文分析了构建纳米（定量）结构—活性关系的研究现状（**第三章**）。所分析的内容包括已有研究所用的数据，已构建的预测模型，金属纳米材料的毒性机理解释以及对此研究领域发展的展望。结果显示现阶段纳米（定量）结构—活性关系的研究主要集中于

对不同细胞系和大肠杆菌的毒性效应的预测（定性以及定量预测）。对模型中描述符的分析显示，金属纳米材料表面修饰分子的亲油性和氢键能力显著影响细胞对纳米材料的摄取。某些情况下，金属纳米材料释放的离子以及产生的活性氧物种是纳米材料引发毒性效应的主因。纳米材料的表面化学特性和尺寸也可以影响其生物效应。对已有研究所用数据的分析显示，有关纳米（定量）结构—活性关系的研究很大程度上仍受限于现有数据的数量和质量。纳米材料的结构表征以及其在溶剂中的结构转化是此类研究中需要注意的问题。

基于上一章对研究现状的分析，**第四章**的研究基于所收集数据构建了纳米结构—活性关系模型。所建模型可对金属纳米材料的危险性进行归类。研究中案例一（半数致死浓度数据，训练集 320 种纳米材料，预测集 80 种纳米材料）和案例三（最小抑菌浓度数据，训练集 133 种纳米材料，预测集 33 种纳米材料）构建了涵盖不同生物类别的综合分类模型。模型的分类预测准确率均在 70%以上。此部分研究同时构建了金属纳米材料对单一生物（斑马鱼，水蚤，羊角月牙藻以及金黄色葡萄球菌）危险性的分类预测模型。所建模型均具有较高的预测准确性。模型所用描述符显示，金属纳米材料的分子极化率、可接近表面积和溶解度对其毒性有显著影响。本章研究有助于纳米材料的危险性分类工作以及相关的环境风险评价。

根据欧洲化学品管理局的要求，如某一纳米材料被归类为其所列出的危险类别之一，需对该纳米材料进行环境风险评价。其中之一是需要估算纳米材料的生态风险阈值和相关的暴露水平。环境暴露水平与生态风险阀值的比值可以表征纳米材料的环境风险。生态风险阈值可以通过计算物种敏感性分布曲线 5%处的数值（通常称为 5%危害浓度，HC5）得到。因此，**第五章**着重于研究金属纳米材料的物种敏感性分布，并把不同的纳米材料结构特征，试验条件以及毒性测试终点考虑在内。基于已有数据，研究得到以不同条件区分的（纳米材料表面包覆、尺寸、形态和毒性暴露时间）纳米银的物种敏感性分布。研究也得到了基于不同毒性测试终点的物种敏感性分布。研究结果显示，聚乙烯吡咯烷酮和柠檬酸钠包覆金属纳米银表面可增强其毒性。不同尺寸纳米银的生物效应只在浓度高时有所差异。不同形态纳米银的物种敏感性分布，以及暴露于纳米

银悬浮液不同时间下的物种敏感性分布之间并无明显差异。结果还表明甲壳类动物对金属纳米材料的敏感性最高。

综上所述，本论文探讨了应用计算毒理学方法评估研究纳米材料安全性的问题，以辅助纳米材料的安全管理。本论文的研究分别着重于对已有金属纳米材料毒性信息的收集和整理，对纳米（定量）结构—活性关系研究现状的分析，以及纳米结构—活性关系和金属纳米材料物种敏感性分布的构建。不可否认，因所用数据数量和质量的问题，本论文所阐述的研究结果仍具有一定的不确定性。这些研究旨在协助纳米材料安全性的评估工作，使决策者获得必要或更充足的信息。为促进该领域的进一步发展，相关后续试验应严格遵照有关准则测试和报道纳米材料的行为归趋和生物效应。可靠以及充足的试验数据对相关研究至关重要。

Acknowledgements

It has been a long journey, one which would not have been possible without the support and help of a great many people. I would firstly like to thank my supervisors Willie and Martina for their patient guidance, encouragement, inspiration, and wise advice. They made my PhD life a very enjoyable and cheerful experience. Thanks also to Arnold and Peter for their time and expertise in reviewing my PhD thesis. Thanks the China Scholarship Council for their financial support of my PhD study in the Netherlands.

I would also like to express my gratitude to all my lovely fellow CML colleagues, including those who already have 'Dr.' in front of their last names and those who are still challenging themselves to achieve the goal. Special thanks to Esther, Jory, Joyce, Maarten, Paul, and Susanna for always being there for us and willing to help us.

I have been lucky enough to have many people standing by my side. I knew Fuyu, Yinlong, and Zhan before coming to the Netherlands; I met Dustin, Enrique, Yudan, Yun, and Zhenyu at Smaragdlaan; I joined Dutch class together with Jiali, Xiaoyu, and Yifen which soon basically turned into a weekly meal out. Having Emmy, Minghou, Pieter Jelle, Stefan, and Verion who have always been supporting me unconditionally was outstanding; I also spent great times with my Leiden squad Eline, Erik, Merijn, and Naomi. Thanks to all these people and also my many other friends with whom I shared unforgettable memories.

Lastly, I would like to give a special thanks to Jiri and everyone from the big Jonkers family. Thanks to my family in China as well, especially my parents. 谢谢爸爸妈妈，我爱你们.

Curriculum Vitae

Guangchao Chen was born in 1987 in Qiqihar, a city in the northeastern China. He attended Qiqihar Experimental Middle School in 2003. From 2006 to 2010, he studied Environmental Sciences at Dalian University, China with a focus in environmental chemistry. He continued to pursue a master's degree in Environmental Sciences from 2010 to 2013 at Dalian University. The project aimed to develop *in silico* models to predict the biodegradability of organic chemicals, which brought him into the field of computational toxicology. In 2013, he was awarded a scholarship by China Scholarship Council (CSC) to follow the Ph.D. program in the Institute of Environmental Sciences (CML), Leiden University. The Ph.D. project was to assist in the risk assessment of engineered nanomaterials by applying computational toxicology into the hazard evaluation of nanomaterials. During this period, he also supervised students in courses and gave lectures to B.Sc. students about nanotechnology and nanotoxicology.

Publications during PhD period

Chen G, Peijnenburg WJ, Xiao Y, Vijver MG. Current Knowledge on the Use of Computational Toxicology in Hazard Assessment of Metallic Engineered Nanomaterials. *International Journal of Molecular Sciences*, 2017, 18(7):1504.

Chen G, Peijnenburg WJ, Xiao Y, Vijver MG. Developing species sensitivity distributions for metallic nanomaterials considering the characteristics of nanomaterials, experimental conditions, and different types of endpoints. *Food and Chemical Toxicology*, 2017. doi: 10.1016/j.fct.2017.04.003.

Xiao Y, Peijnenburg WJ, **Chen G**, Vijver MG. Toxicity of copper nanoparticles to Daphnia magna under different exposure conditions. *Science of the Total Environment*. 2016, 563-564:81-8.

Hua J, Vijver MG, **Chen G**, Richardson MK, Peijnenburg WJ. Dose metrics assessment for differently shaped and sized metal-based nanoparticles. *Environmental Toxicology and Chemistry*. 2016, 35(10):2466-2473.

Chen G, Peijnenburg WJ, Kovalishyn V, Vijver MG. Development of nanostructure–activity relationships assisting the nanomaterial hazard

categorization for risk assessment and regulatory decision-making. *RSC Advances*. 2016, 6:52227-52235.

Chen G, Vijver MG, Peijnenburg WJ. Summary and analysis of the currently existing literature data on metal-based nanoparticles published for selected aquatic organisms: Applicability for toxicity prediction by (Q)SARs. *Alternatives to Laboratory Animals*. 2015, 43(4):221-40.

Xiao Y, Vijver MG, **Chen G**, Peijnenburg WJ. Toxicity and accumulation of Cu and ZnO nanoparticles in Daphnia magna. *Environmental Science & Technology*. 2015, 49(7):4657-64.

Chen G, Li X, Chen J, Zhang YN, Peijnenburg WJ. Comparative study of biodegradability prediction of chemicals using decision trees, functional trees, and logistic regression. *Environmental Toxicology and Chemistry*. 2014, 33(12):2688-93.

Submitted work

Chen G, Vijver MG, Xiao Y, Peijnenburg WJ. Recent advances towards the development of (quantitative) structure-activity relationships for metallic nanomaterials: A critical review. *Materials*. Under revision.

Peijnenburg WJ, **Chen G**, Vijver MG. Nano-QSAR for environmntal hazard assessment: turning challenges into opportunities. A. Gajewicz and T. Puzyn (Eds.). Computational Nanotoxicology: Challenges, pitfalls and perspectives. Pan Stanford Publishing. 2017. Submitted.

Xiao Y, Peijnenburg WJ, **Chen G**, Vijver MG. Impact of water chemistry on the particle-specific toxicity of copper nanoparticles to Daphnia magna. *Science of the Total Environment*. Under revision.