
Article:
Tantra, R, Oksel, C, Puzyn, T et al. (5 more authors) (2015) Nano(Q)SAR: Challenges, pitfalls and perspectives. Nanotoxicology, 9 (5). pp. 636-642. ISSN 1743-5390

https://doi.org/10.3109/17435390.2014.952698

Nano-(Q)SAR: Challenges, pitfalls and perspectives

Ratna Tantra‡, Ceyda Oksel*, Tomasz Puzyn†, Jian Wang‡, Kenneth Robinson‡, Xue Z. Wang*, Cai Y. Ma* and Terry Wilkins*

*Institute of Particle Science and Engineering, School of Process, Environmental and Materials Engineering, University of Leeds, Leeds, LS2 9JT, UK

† Laboratory of Environmental Chemometrics, Faculty of Chemistry, University of Gdansk, Wita Stwosza 63, 80-952 Gdansk, Poland

‡ National Physical Laboratory, Hampton Road, Teddington, Middlesex, TW11 0LW, UK

Corresponding author: ratna.tantra@npl.co.uk

Abstract

Regulation for nanomaterial is urgently needed and the drive to adopt an intelligent testing strategy is evident. The intelligent testing strategy will not only be beneficial from a cost reduction point of view but will also mean the reduction of the moral and ethical concerns related to animal research. In the chemical and legislative world, such an approach is promoted by REACH and in particular the use of (Q)SAR as a tool for the purpose of categorisation. In addition to traditional compounds, (Q)SAR has also been applied to nanomaterials i.e. nano(Q)SAR, useful to correlate toxicological endpoints with physicochemical properties. Although (Q)SAR in chemicals is well established, nano(Q)SAR is still at an early stage of its development and its successful uptake is far from reality. The purpose of this paper is to identify some of the pitfalls and challenges associated with nano-(Q)SARs, in relation for its use to categorise nanomaterials. Our findings show clear gaps in the research framework that must be addressed if we are to have reliable predications from the use of such models. Three major types of barriers were identified: a) the need to improve quality of experimental data in which the models are being developed from in the first place, b) the need to have practical guidelines for the development of the nano(Q)SAR models, c) the need to standardise and harmonise activities for the purpose of regulation. Out of the three barriers, immediate attention is needed for a) as this underpins activities associated in b) and c). It should be noted that the usefulness of data in the context of nano-(Q)SAR modelling is not only about the quantity of data but also about the quality, consistency and accessibility of those data.

Introduction

The National Science Foundation has estimated that nanotechnology is to be worth $1 trillion by 2015 [1] and that engineered nanomaterials represent a growing class of material being introduced into multiple business sectors. However, there are also concerns surrounding the potential harmful effects of engineered nanomaterial imposed on health and the environment. The immediate goal to regulate without hampering public perception on the benefits of nano-enabled products but scientific findings have not yet provided any clear answers on the toxicity of nanomaterials [2]. The real danger here is that regulation will run ahead of scientific reality, which may lead on the tightening or loosening of regulations, thus affecting either the nanomaterial manufacture sectors or the environment and human health.

The need to develop a sound infrastructure for assessment of risk is clear. However, the variety of nanomaterials is great. Subsequently, this will mean the need to test large number of nanomaterial/matrix combination; this makes it difficult to assess the possible hazard and

risk of each nanomaterial. In the chemical world, an "intelligent testing" approach is employed in order to improve testing efficiency. One strategy to alleviate the problem is through the establishment of chemical categories, which relies on "similarity principle" i.e. similarities in the structure of different molecules/compounds that may create predictable patterns of particular biological activity/toxicity endpoints. This allows untested chemicals to be categorised by assuming that the biological activity is linked to their molecular structure, as determined by compositional and structural descriptors such as size, elemental composition, topology and functional groups [3]. Categorisation has meant the ability to streamline hazard and risk assessment process; where costs are concerned, this will mean that a particular chemical group can be assessed as a whole, without the need to test every category member for every regulatory endpoint. Moreover, in relation to ethical considerations, a great deal of animal experimentation will be reduced. This is particularly attractive in certain market categories, such as European cosmetic products (e.g. skin creams, oral care, shampoo, conditioner, deodorants, etc.) in which restrictions on in vivo testing exist.

The development of chemical categories often involves the need to estimate missing entries. This can be achieved through the use of trend analysis, which often involves a form of interpolation and extrapolation. Interpolation can be defined as the estimation of a value between two known data points [4,5]. Extrapolation is an estimate of a value based on extending a known sequence of values near or at the category boundary using measured values from internal category members [5]. Another approach to estimate missing experimental values is to use read across, which involves using information on one chemical structure and making some assessment about the relevance of that information for a second chemical structure considered similar [5]. Hence, methods of extrapolation and read across are therefore inherently more risky than interpolation. In addition to trend analysis and read across, the use of in silico methods such as (Q)SAR, [6] is attractive to many industry and regulators. The predictive ability of in silico methods resulted in REACH to promote the use of (Q)SAR to develop chemical categories [7,8].

It is not the intent of this paper to delve into the details of the processes involved in (Q)SAR, as they have been covered in great details elsewhere and the reader is referred to the relevant past literature [7,9-11]. Only a brief summary will be given in order to familiarise the reader with the basics of this topic. (Q)SAR is a collective term that refers to both Structure-Activity Relationships (SAR) and Quantitative Structure-Activity Relationships (QSAR). Both are theoretical models, used to predict "analysed activity" (e.g. some biological endpoints) from "descriptor values". Whereas QSAR is a numerical measure i.e. mathematical (regression), SAR is more qualitative in nature i.e. categorisation. In relation to chemicals the descriptors are often referred to as molecular descriptors, which can range from geometrical and topological (e.g. moment of inertia, accessible surface area and volume, aspect ratios) to electronic (e.g. HOMO and LUMO energies, dipole moment); the descriptors here can either be measured (experimental) or calculated (theoretical numerical values). (Q)SAR model is created using a data set referred to as the training set e.g. chemicals for which activity is known**.** In QSAR, the model is constructed by using classical chemoinformatic methods, such as classical linear methods (e.g. partial least squares and multiple linear regression), and non-linear methods (e.g. support vector machine and artificial neural networks) [12-16]. Modellers also sometimes employ principal component analysis (PCA) to aid the process of designing (Q)SARs [17,18] e.g. for searching for structural similarity patterns and thus predefining new categorises of the studied chemicals. In SAR, as there is no attempt to derive a quantitative model, artificial intelligence methods such as decision trees and discriminant analysis are used instead[19]. The most crucial steps in (Q)SAR are to demonstrate the

robustness of the model (validation) and to set the boundaries of the validated model which define the domain of applicability. This helps to check which external chemicals can be predicted using the built model and to establish how the model will perform when faced with compounds that were not included in the training or test set.

The use of (Q)SARs in order to categorise chemicals is quite advanced, which in turn have allowed past workers to create a broad hazard identification profile, prioritise chemicals and estimate values for untested chemicals[20-28]. Furthermore, the development of chemical categories have also been used to identify hazards (associated with safe storage, handling and disposal of waste)[29-32], prediction of physicochemical properties or toxic/biochemical effects[33-36] and provide information about mechanism of action[37,38], thus promoting the safe use of chemicals.

Unlike chemicals, no clear strategy is in place to categorise nanomaterials. Several suggestions have been made. Glotzer and Solomon[39] suggests that nanomaterials can be categorised through the use of eight orthogonal dimensions (surface coverage, aspect ratio, faceting, pattern quantization, branching, chemical ordering, shape gradient, and roughness) in order to describe the key attributes of nanomaterials. Hansen and Stone suggested the use of a chemistry based categorisation system as a starting point[40,41]. Although little work has been done in relation to the use of (Q)SAR to categorise nanomaterial, the approach (of using such models much in the same way as in chemicals) would be attractive, if proven possible.

In 2009, Puzyn coined the term nano(Q)SAR i.e. (Q)SAR models that correlate the physicochemical properties of nanomaterials to their biological activity[42]. To date, several alternative acronyms on top of nano-(Q)SAR have been used to describe the same method, which includes: Quantitative Nanostructure-Activity Relationships (QNAR)[43] and Quantitative Nanostructure-Toxicity Relationships (QNTR)[44]. For the remainder of the manuscript the term of nano(Q)SAR shall be used throughout. Unlike the use of (Q)SAR in chemicals, nano-(Q)SAR is still at an early stage[45], far from successful uptake in relation to nano-regulation.

The purpose of the paper is to understand the pitfalls and challenges associated with nano(Q)SAR and research requirements in order to develop reliable nanomaterial categories. In order to meet this objective, we will be reviewing past literature on nano-(Q)SAR and report relevant findings to their use for categorisation of nanomaterials. We will present criteria necessary in the use of (Q)SAR to categorise chemicals and translate this to nano(Q)SAR. Finally, we will identify the pitfalls/barriers associated and discuss future work necessary for implementation.

Much of the source material presented for this paper is from peer-reviewed papers as well as grey literature e.g. guidelines/scientific opinion documents for regulations, with relevance to the objectives stated.

Results and Discussion

Nano(Q)SARs: recent advances

To date, only a handful of studies on nano(Q)SAR have been carried out, with most studies reporting metal oxide nanoparticles due to their relatively high volume of use/production. . Different modellers have attempted to establish a link between different physicochemical descriptors and observed biological activity. Table 1 summarises some of the different predictors used by different researchers. A total of twelve descriptors have been associated

with nano(Q)SAR; these are the physicochemical properties of potential relevance to toxicity, although no clear and valid mechanistic understanding has been developed. There are some common descriptors i.e. have been reported by more than once. These are: particle size, zeta-potential, surface modifications and spin-spin relaxivities. Out of all of these, surface modification is the most common, indicating the importance of this property and its relevance to toxicity.

Table 1 Past nano(Q)SAR studies and the use of different descriptors of relevance. Here, ΔH_{Me+} represents the enthalpy of formation of a gaseous cation having the same oxidation state as that in the metal oxide structure.

Type of model/ (References)	**Particle Size**	**Zeta-potential**	**Surface Modifications e.g. surface coatings**	**Spin-spin relaxivities**	**Others**
nanoQSAR [46]	√	√	√	√	X
nanoQSAR [47]	X	X	X	X	ΔH_{Me+}
nanoQSAR [10,48]	X	X	X	X	Band energy
nanoQSAR [46]	X	X	√	X	X
nanoSAR [49]	√	X	X	X	atomisation energy and nanoparticle volume fraction (in solution)
nanoQSAR [50]	X	X	√	X	X
nanoQSAR [51]	X	X	√	X	hydrogen-bond donor sites
nanoQSAR [52]	X	X	√	X	X
nanoSAR [53,54]	√	√	X	√	conduction band energy, ionic index, spin-lattice
Total number of ticks	3	2	5	2	

From the table, it is clear that apart from Liu and co-workers (who have developed nanoSAR) [49,53,54], most of the studies have been associated with nanoQSAR. In general, much work is still needed in the field, in particular the application of nano(Q)SAR to the development of nanomaterial categories. The first step is to identify the criteria imposed for uptake, which will be similar to those identified for corresponding chemicals. According to an OECD guideline[7] on (Q)SAR, the basic criteria that must be fulfilled is the generation of reliable and validated models, which will allow confidence in the predictions made. However, barriers exists, which prevent the criteria from being fulfilled and these will be further assessed below. Once these basic criteria are fulfilled, then there is a need to have a clear implementation route i.e. towards industry uptake and regulatory acceptance.

Barrier 1: the generation of a reliable model

Unlike chemicals, measuring the physicochemical characteristics of nanomaterials is not straightforward with current instrumentations. From a scientific perspective, nanomaterials cannot be considered as a homogeneous group and subsequently this means that getting reliable data is not easy to achieve. Potentially, this leads to a situation in which experimental data gets reported without proper understanding of the associated errors and the propagation of these errors through the model. Sources of errors may arise from a number of factors including polydispersity, biological environment and inappropriate measurement.

In relation to the polydispersity of nanomaterial, it has been argued by Baalousha and Lead[55] that most nanomaterials tested are too polydisperse. Materials close to monodispersity are needed in order to have better reliability of result findings associated with studying environmental behaviour, dose, structure–activity relationships and mechanisms of toxicity. Although there is great effort in the scientific community to develop test/reference materials e.g. OECD Working Party on Manufactured Nanomaterials[56]), there is a need to address whether the commercially relevant nanomaterials offered held in this repository are suitable for the purpose nano(Q)SAR models. If monodisperse and homogeneous nanomaterial sample is needed, then only a handful of nanomaterials in existence have been certified and sold under the banner of reference nanomaterials, to include National Institute of Standards and Technology (NIST) gold nanoparticle reference materials (10, 30 and 60 nm). The use of such materials are ideal for use in nanotoxicology studies and recently they been shown suitable to act as negative controls for nanoparticle genotoxicity studies[57].

With respect to the analytical techniques currently available, it is difficult to measure accurately, a highly polydisperse sample. In a recent study, Anderson et al.[58] show that complex particle size distributions i.e. away from the simple monomodal distribution will result in large data variability. It was reported that light scattering based Particle Tracking Analysis and Dynamic Light Scattering platforms were only able to detect a single population of particles corresponding either the largest or smallest particles in a multimodal sample. Clearly, the inadequacy of the instrumental methods to characterise nanomaterials is a huge barrier in this field, as echoed by several past workers[2,55].

In addition to polydispersity issue, the complex biological matrix in which the particles are dispersed in can also pose problems where measurement is concerned, potentially resulting in agglomeration and thus the formation an unstable suspension. Furthermore, nanomaterial-media interactions can be dynamic in nature, which may pose further difficulties for the instrument to measure under such conditions. Finally, physicochemical properties measured may not being directly associated with the observed biological effects. Due to the analytical challenges posed in a) and b), some studies have characterised nanomaterials in their “pristine” state i.e. absence of the actual biological test media. In fact, few studies have assessed the potential transformation of nanomaterials in an environmental or mammalian system[59,60].

In addition to measurement of physicochemical properties, measurements of biological endpoints may also be problematic. In in vitro measurements, there are several biological endpoints that reflect changes associated with cell activity, which may be employed to indicate toxicity hazard related to the nanomaterial under investigation. Examples include: evidence of appreciable cell death relative to suitable control experiments, growth retardation and cell membrane damage. There are several bioassay tests that can measure these biological

endpoints but the results may not be reliable. The following potential sources of errors, which may lead to false interpretation, have been previously identified as being problematic:

a) endotoxin (LPS, lipopolysaccharide) contamination in the nanomaterial[61].

b) choice of inappropriate end-points for the nanomaterial. There is a need to define a standard set of biological assays (and protocols) clearly in order to evaluate the overall in vitro (and in vivo) response of the tested nanoparticles. The assay chosen should be truly indicative of key activity, property or toxicological effects caused by these nanomaterials[62].

c) interference in the assay readout by the nanomaterials e.g. tetrazolium based assays and the subsequent potential interference of the nanomaterials with the formazan salts[63,64].

d) varied nanomaterial dispersion protocols, including differences in the amount, source and pre-treatment of serum proteins used (affecting particle size distribution and agglomerates/aggregates population, etc.)[65,66].

The problems identified so far imply the need to develop better research techniques/ methods and that the developed methods are validated. Eurachem[67] clearly states that analytical measurements should be made using methods and equipment which have been tested to ensure that they are fit for purpose. The current state of toxicological research in nanomaterials is that methods developed are rarely validated.

Once the conditions of the method validation are met, only then can researchers consider a higher metrological standard of measurement, conduct uncertainty analyses to estimate uncertainty and the propagation of uncertainty. The need to have reliable data for the nano-(Q)SAR model may mean that for the first step of developing a robust nano-(Q)SAR model, it would be easier (more achievable) to build a model for engineered nanomaterials using homogeneous, monodisperse samples in environments where agglomeration of the nanomaterials would be relatively low.

Another factor that influences the reliability of the nano(Q)SAR model is in the selection of descriptors. Having an agreement on suitable descriptors is important, as past workers found that the choice of descriptors can affect the quality, significance and interpretation of nano-(Q)SAR model[10,52]. It may be that descriptors can be selected or adapted from traditional chemical descriptors, but it is highly likely that descriptors specifically developed for nanomaterials are needed. For example, Borders et al have shown that new types of defects in carbon nanotubes can be represented as a new descriptor in the prediction of its mechanical properties[68]. In a recent study, Wang et al.[69] have shown how Principal Component Analysis (PCA) can be used in order to identify potentially suitable descriptors. Through the use of PCA, they identified: particle charge, aspect ratio and metal content as potential descriptors. In summary, much attention must be given to the selection of trial descriptor sets when developing nano(Q)SAR models, as they play a pivotal role in predictive quality of the model.

Barrier 2: validation of the model

A pre-requisite for the uptake of (Q)SAR[7] and subsequently nano(Q)SAR, is for the model to be validated. It states clearly that the principles for validation should include: a defined endpoint, an unambiguous algorithm, a defined domain of applicability, appropriate measures of goodness-of-fit, robustness, predictivity and a mechanistic interpretation, if possible. Validation is vital to ensure that the predictive ability of the model is not due to chance factors.

Validation for nano(Q)SAR is not straightforward due to the following reasons:

a) The heterogeneity of nanomaterial family: There is a need to develop separate models that are specific to nanomaterial types and thus properties[70]. Due to large number of nanomaterial types that can be engineered (and subsequently different mechanisms of toxicity), it has been suggested by Puzyn that individual classes of nanomaterials should be modelled separately [42].

b) The paucity of data available for nano(Q)SAR: Since nano(Q)SARs are developed with statistical methods (MLR, neural networks etc.), the variance in the group of nanomaterial needs to be sufficiently represented in the training and validation sets of particles. It is commonly accepted that the ratio between the number of descriptors and compounds in the training set should be at least as 1 to 5[71,72].

c) The lack of standardized validation metrics: The approach used for validation needs to be defined, for example whether to employ variants of cross validation or external test set validation[73]. Ideally, validation should be done externally i.e. by an external predication set, in which the test data set has to be independent not only from model building but also from model selection[74]. However, this step may not be always possible. If additional testing is not feasible then an internal validation may be carried out e.g. using the "leave one out" cross validation (CV) method; CV is carried out by omitting a point and then calculating the value of this location using the remaining points[75]. Having said this, Tropsha et al.[76] demonstrated that leave-one-out cross validation and external test set metrics do not correlate[77] and stated that a *"high value* of leave-one-out cross-validated R^2 appears to be a necessary but not sufficient condition for a model to have a high predictive power". It should also be noted that the model's predictivity can only be confirmed with external validation without which the (Q)SAR modelling procedure is incomplete.

Barrier 3: improving industry uptake and regulatory acceptance

In relation to chemicals, there seems to be a clear regulatory application of (Q)SAR in the US but this is less so in Europe[78]. The wider acceptance of (Q)SAR in the US is attributed to the need to evaluate chemical substances within a relatively short period of time. The US Environmental Protection Agency (EPA) in particular is keen to promote the use of (Q)SAR to develop chemical category in hazard and risk assessment[79]. In the European Union, regulatory acceptance of (Q)SARs is limited. There is a push by REACH, through activities arising from the existing OECD chemicals programme[7,80], to promote the use of (Q)SAR as an alternative method to evaluate chemicals. Clearly, the fact that implementation of (Q)SAR in Europe is still problematic for chemicals, means that nano(Q)SARS will not easily be accepted in the near future. In order to implement nano(Q)SAR it is vital to demonstrate to regulators, and industry, that nano(Q)SARs are scientifically validated and clear explanations on how to use such models for making decisions[45] should also be given. Once this is achieved, we can then "harmonise activities" e.g. by forging internationally agreed document standards and guidelines. Guidelines of relevance should include the provision of detailed guidance in relation to the practicalities on the use of nano-(Q)SAR e.g. detailing how to identify acceptability criteria, how to generate adequate and relevant descriptors[81-83].

Conclusion

There is a widespread regulatory and scientific interest in developing intelligent and cost effective testing. In particular REACH is promoting the use of alternative methods such as (Q)SAR for the purpose of categorising chemicals. One of the most important characteristics of (Q)SAR is in their predictive power, having been seen as "an enabler" in bringing new chemicals to commercialisation.

The recent application of (Q)SAR to nanomaterials shows that researchers are starting to use such models in order to categorise nanomaterials in the same way as chemicals. This may be beneficial as it aims to improve the efficiency of hazard and risk assessment. This may be is useful in market categories where there is a restriction on in vivo testing. However, it is clear from this review that the valid implementation of nano-(Q)SAR is a long way off, as past experience associated even with chemicals shows the considerable amount of time and effort needed to implement the use of such tools at an internationally acceptable level.

Several issues have been highlighted in this review, which cast serious doubts over the reliability of such models to support nano-regulation. If nano-(Q)SAR is to be implemented, then issues associated with experimental data quality used to develop the model in the first place must be tackled. We have identified the need for better analytical techniques to deal with polydispersity in a sample and when the nanomaterial is dispersed in complex media. In addition, there is a need to have validated methods and more ideal test/reference materials. These activities are valuable for any understanding of nanomaterials not just for nano(Q)SARS. In relation to the development of the model itself, there is a need to generate practical guidance e.g. identification of relevant descriptors for nanomaterials. Once the issues associated with reliability have been tackled, the next step is to ensure better co-ordination between the scientific community with industry and regulatory authorities.

Overall, the first step is to generate reliable nano(Q)SAR models and that as part of the validation process, an external validation with independent series of data should be used. If this can be achieved, internationally agreed documentary standards and guidelines can be generated.

Acknowledgements

The authors would like to thank the European Commission for financial support via the FP7 NanoReg and DEFRA UK (Department for Environment, Food and Rural Affairs). TP is grateful for the financial support of the European Commission through Marie Curie IRSES program, NanoBRIDGES project (FP7-PEOPLE-2011-IRSES, Grant Agreement Number 295128), COST action TD1204: Modelling Nanomaterial Toxicity (MODENA) and Foundation for the Polish Science (FOCUS Program). Lastly, many thanks to Stuart Hewlin (Procter and Gamble, UK) for his valuable assistance during the drafting of this paper.

References

(1) Roco, M. C.; Bainbridge, W. S.: Societal implications of nanoscience and nanotechnology: Maximizing human benefit. J. Nanopart. Res. **2005**, 7, 1-13.

(2) Tantra, R.; Shard, A.: We need answers. Nat. Nanotechnol. **2013**, 8, 71-71.

(3) Bajorath, J.: Chemoinformatics: concepts, methods, and tools for drug discovery; Springer, 2004; Vol. 275.

(4) Duffus, J. H.; Nordberg, M.; Templeton, D. M.: Glossary of terms used in toxicology, (IUPAC recommendations 2007). Pure Appl. Chem. **2007**, 79, 1153-1344.

(5) Patlewicz, G. "Chemical Categories and Read Across," Joint Research Centre, 2005.

(6) Lahl, U.; Gundert-Remy, U.: The use of (Q) SAR methods in the context of REACH. Toxicol. Mech. Methods **2008**, 18, 149-158.

(7) OECD: Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship Models. 2007.

(8) OECD: Guidance on the Grouping of Chemicals. **2007**.

(9) van Leeuwen, C.; Vermeire, T.; Vermeire, T.: Risk assessment of chemicals: an introduction; Springer, 2007.

(10) Burello, E.; Worth, A. P.: QSAR modeling of nanomaterials. Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology **2011**, 3, 298-306.

(11) Dudek, A. Z.; Arodz, T.; Galvez, J.: Computational methods in developing quantitative structure-activity relationships (QSAR): a review. Comb. Chem. High Throughput Screening **2006**, 9, 213-228.

(12) Wessel, M. D.; Jurs, P. C.: Prediction of reduced ion mobility constants from structural information using multiple linear regression analysis and computational neural networks. Anal. Chem. **1994**, 66, 2480-2487.

(13) Fischer, M.: Computational neural networks: a new paradigm for spatial analysis. Environment and Planning A **1997**, 30, 1873-1891.

(14) Guha, R.; Jurs, P. C.: Interpreting computational neural network QSAR models: a measure of descriptor importance. J. Chem. Inf. Model. **2005**, 45, 800-806.

(15) Yao, X.; Panaye, A.; Doucet, J.-P.; Zhang, R.; Chen, H.; Liu, M.; Hu, Z.; Fan, B. T.: Comparative study of QSAR/QSPR correlations using support vector machines, radial basis function neural networks, and multiple linear regression. J. Chem. Inf. Comput. Sci. **2004**, 44, 1257-1266.

(16) Winkler, D. A.; Burden, F. R.: Application of Neural Networks to Large Dataset QSAR, Virtual Screening, and Library Design #. In T Combinatorial Library, 2002; Vol. 201; pp 325-367.

(17) Eriksson, L.; Andersson, P. L.; Johansson, E.; Tysklind, M.: Megavariate analysis of environmental QSAR data. Part I–A basic framework founded on principal component analysis (PCA), partial least squares (PLS), and statistical molecular design (SMD). Mol. Diversity **2006**, 10, 169-186.

(18) Cardin, R.; Michielan, L.; Moro, S.; Sperduti, A.: PCA-Based Representations of Graphs for Prediction in QSAR Studies. In Artificial Neural Networks–ICANN 2009; Springer, 2009; pp 105-114.

(19) Sussman, N.; Arena, V.; Yu, S.; Mazumdar, S.; Thampatty, B.: Decision tree SAR models for developmental toxicity based on an FDA/TERIS database. SAR QSAR Environ. Res. **2003**, 14, 83-96.

(20) Verhaar, H.; Van Leeuwen, C.; Bol, J.; Hermens, J.: Application of QSARs in risk management of existing chemicals. SAR QSAR Environ. Res. **1994**, 2, 39-58.

(21) Hulzebos, E.; Janssen, P.; Maslankiewicz, L.; Meijerink, M.; Muller, J.; Pelgrom, S.; Verdam, L.; Vermeire, T.: The application of structure-activity relationships in human hazard assessment: a first approach. **2001**.

(22) Cronin, M. T.; Jaworska, J. S.; Walker, J. D.; Comber, M. H.; Watts, C. D.; Worth, A. P.: Use of QSARs in international decision-making frameworks to predict health effects of chemical substances. Environ. Health Perspect. **2003**, 111, 1391.

(23) Patlewicz, G.; Ball, N.; Booth, E. D.; Hulzebos, E.; Zvinavashe, E.; Hennes, C.: Use of Category approaches, Read-across and (Q) SAR: General considerations. Regulatory Toxicology and Pharmacology **2013**.

(24) Tebby, C.; Mombelli, E.; Pandard, P.; Péry, A. R.: Exploring an ecotoxicity database with the OECD (Q) SAR Toolbox and DRAGON descriptors in order to prioritise testing on algae, daphnids, and fish. Sci. Total Environ. **2011**, 409, 3334-3343.

(25) Petkov, P.; Temelkov, S.; Villeneuve, D.; Ankley, G.; Mekenyan, O.: Mechanism-based categorization of aromatase inhibitors: a potential discovery and screening tool. SAR QSAR Environ. Res. **2009**, 20, 657-678.

(26) Bolcic-Tavcar, M.; Vracko, M.: Assessing the reproductive toxicity of some (con) azole compounds using a structure-activity relationship approach. SAR QSAR Environ. Res. **2009**, 20, 7-8.

(27) Gubbels-Van Hal, I.: Non-testing in REACH? Chimica oggi **2007**, 25, 14-16.

(28) Mekenyan, O.; Dimitrov, S.; Pavlov, T.; Veith, G.: POPs: A QSAR system for developing categories for persistent, bioacculative and toxic chemicals and their metabolites. SAR QSAR Environ. Res. **2005**, 16, 103-133.

(29) Byers, D.; Szerlog, M.; Air; Waste Management, A.: A rapid and low cost hazard categorization technique for identifying hazardous wastes, 1997.

(30) Mansouri, K.; Ringsted, T.; Ballabio, D.; Todeschini, R.; Consonni, V.: Quantitative Structure–Activity Relationship Models for Ready Biodegradability of Chemicals. J. Chem. Inf. Model. **2013**, 53, 867-878.

(31) Lowe, M. F.; McCall, D.: Globally Harmonized System of Classification and Labelling of Chemicals (GHS). US Environmental Agency, Arlington USA **2009**.

(32) Cronin, M. T.; Walker, J. D.; Jaworska, J. S.; Comber, M. H.; Watts, C. D.; Worth, A. P.: Use of QSARs in international decision-making frameworks to predict ecologic effects and environmental fate of chemical substances. Environ. Health Perspect. **2003**, 111, 1376.

(33) Goldberg, L.: Structure-activity correlation as a predictive tool in toxicology. HEMISPHERE PUBL. CORP., 19 WEST 44 TH ST., NEW YORK, NY 10036, USA. 1983. **1983**.

(34) Cassani, S.; Kovarich, S.; Papa, E.; Roy, P. P.; van der Wal, L.; Gramatica, P.: Daphnia and fish toxicity of (benzo) triazoles: Validated QSAR models, and interspecies quantitative activity-activity modelling. J. Hazard. Mater. **2013**.

(35) Mayer, J.; Cheeseman, M.; Twaroski, M.: Structure–activity relationship analysis tools: validation and applicability in predicting carcinogens. Regulatory Toxicology and Pharmacology **2008**, 50, 50-58.

(36) McKinney, J. D.; Richard, A.; Waller, C.; Newman, M. C.; Gerberick, F.: The practice of structure activity relationships (SAR) in toxicology. Toxicol. Sci. **2000**, 56, 8-17.

(37) Benigni, R.; Bossa, C.: Predictivity and reliability of QSAR models: the case of mutagens and carcinogens. Toxicol. Mech. Methods **2008**, 18, 137-147.

(38) Liszeková, D.; Polakovičová, M.; Beňo, M.; Farkaš, R.: Molecular Determinants of Juvenile Hormone Action as Revealed by 3D QSAR Analysis in Drosophila. PLoS One **2009**, 4, e6001.

(39) Glotzer, S. C.; Solomon, M. J.: Anisotropy of building blocks and their assembly into complex structures. Nat. Mater. **2007**, 6, 557-562.

(40) Foss Hansen, S.; Larsen, B. H.; Olsen, S. I.; Baun, A.: Categorization framework to aid hazard identification of nanomaterials. Nanotoxicology **2007**, 1, 243-250.

(41) Stone, V.; Nowack, B.; Baun, A.; van den Brink, N.; von der Kammer, F.; Dusinska, M.; Handy, R.; Hankin, S.; Hassellöv, M.; Joner, E.: Nanomaterials for environmental studies: classification, reference material issues, and strategies for physico-chemical characterisation. Sci. Total Environ. **2010**, 408, 1745-1754.

(42) Puzyn, T.; Leszczynska, D.; Leszczynski, J.: Toward the Development of "Nano‐QSARs": Advances and Challenges. Small **2009**, 5, 2494-2509.

(43) 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 Screening **2011**, 14, 217-225.

(44) Winkler, D. A.; Mombelli, E.; Pietroiusti, A.; Tran, L.; Worth, A.; Fadeel, B.; McCall, M. J.: Applying quantitative structure-activity relationship approaches to nanotoxicology: current status and future potential. Toxicology **2012**.

(45) Mays, C.; Benfenati, E.; Pardoe, S.: Use and perceived benefits and barriers of QSAR models for REACH: findings from a questionnaire to stakeholders. Chemistry Central Journal **2012**, 6, 1-9.

(46) Fourches, D.; Pu, D.; Tassa, C.; Weissleder, R.; Shaw, S. Y.; Mumper, R. J.; Tropsha, A.: Quantitative Nanostructure−Activity Relationship Modeling. ACS Nano **2010**, 4, 5703-5712.

(47) Puzyn, T.; Rasulev, B.; Gajewicz, A.; Hu, X.; Dasari, T. P.; Michalkova, A.; Hwang, H.-M.; Toropov, A.; Leszczynska, D.; Leszczynski, J.: Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles. Nat. Nanotechnol. **2011**, 6, 175-178.

(48) Burello, E.; Worth, A. P.: A theoretical framework for predicting the oxidative stress potential of oxide nanoparticles. Nanotoxicology **2011**, 5, 228-235.

(49) Liu, R.; Rallo, R.; George, S.; Ji, Z.; Nair, S.; Nel, A. E.; Cohen, Y.: Classification NanoSAR development for cytotoxicity of metal oxide nanoparticles. Small **2011**, 7, 1118-1126.

(50) Epa, V. C.; Burden, F. R.; Tassa, C.; Weissleder, R.; Shaw, S.; Winkler, D. A.: Modeling Biological Activities of Nanoparticles. Nano Lett. **2012**, 12, 5808-5812.

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

(52) Shao, C.-Y.; Chen, S.-Z.; Su, B.-H.; Tseng, Y. J.; Esposito, E. X.; Hopfinger, A. J.: Dependence of QSAR Models on the Selection of Trial Descriptor Sets: A Demonstration Using Nanotoxicity Endpoints of Decorated Nanotubes. J. Chem. Inf. Model. **2013**, 53, 142-158.

(53) 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 **2013**.

(54) Liu, R.; Zhang, H. Y.; Ji, Z. X.; Rallo, R.; Xia, T.; Chang, C. H.; Nel, A.; Cohen, Y.: Development of Structure-Activity Relationship for Metal Oxide Nanoparticles. Nanoscale **2013**.

(55) Baalousha, M.; Lead, J. R.: Nanoparticle dispersity in toxicology. Nat. Nanotechnol. **2013**, 8, 308-309.

(56) Kearns, P.; Gonzalez, M.; Oki, N.; Lee, K.; Rodriguez, F.: The safety of nanotechnologies at the OECD. In Nanomaterials: Risks and Benefits; Springer, 2009; pp 351-358.

(57) Nelson, B. C.; Petersen, E. J.; Marquis, B. J.; Atha, D. H.; Elliott, J. T.; Cleveland, D.; Watson, S. S.; Tseng, I.-H.; Dillon, A.; Theodore, M.: NIST gold nanoparticle reference materials do not induce oxidative DNA damage. Nanotoxicology **2013**, 7, 21-29.

(58) Anderson, W.; Kozak, D.; Coleman, V. A.; Jämting, Å. K.; Trau, M.: A comparative study of submicron particle sizing platforms: Accuracy, precision and resolution analysis of polydisperse particle size distributions. J. Colloid Interface Sci. **2013**, 405, 322-330.

(59) Fadeel, B.; Savolainen, K.: Broaden the discussion. Nat. Nanotechnol. **2013**, 8, 71-71.

(60) McCall, M. J.; Coleman, V. A.; Herrmann, J.; Kirby, J. K.; Gardner, I. R.; Brent, P. J.; Johnson, C. M.: A tiered approach. Nat. Nanotechnol. **2013**, 8, 307-308.

(61) Esch, R. K.; Han, L.; Foarde, K. K.; Ensor, D. S.: Endotoxin contamination of engineered nanomaterials. Nanotoxicology **2010**, 4, 73-83.

(62) Monteiro-Riviere, N. A.; Tran, C. L.: Nanotoxicology: Progress toward Nanomedicine; CRC Press, 2014.

(63) Neumeyer, A.; Bukowski, M.; Veith, M.; Lehr, C.-M.; Daum, N.: Non-invasive determination of cellular oxygen consumption as novel cytotoxicity assay for nanomaterials. Nanotoxicology **2014**, 8, 50-60.

(64) Hsiao, I.-L.; Huang, Y.-J.: Improving the interferences of methyl thiazolyl tetrazolium and IL-8 assays in assessing the cytotoxicity of nanoparticles. J. Nanosci. Nanotechnol. **2011**, 11, 5228-5233.

(65) Murdock, R. C.; Braydich-Stolle, L.; Schrand, A. M.; Schlager, J. J.; Hussain, S. M.: Characterization of nanomaterial dispersion in solution prior to in vitro exposure using dynamic light scattering technique. Toxicol. Sci. **2008**, 101, 239-253.

(66) Johnston, H.; Pojana, G.; Zuin, S.; Jacobsen, N. R.; Møller, P.; Loft, S.; Semmler-Behnke, M.; McGuiness, C.; Balharry, D.; Marcomini, A.: Engineered nanomaterial risk. Lessons learnt from completed nanotoxicology studies: potential solutions to current and future challenges. Critical reviews in toxicology **2013**, 43, 1-20.

(67) Eurachem, G.: The fitness for purpose of analytical methods. A laboratory guide to method validation and related topics **1998**.

(68) Borders, T. L.; Fonseca, A. F.; Zhang, H.; Cho, K.; Rusinko III, A.: Developing Descriptors To Predict Mechanical Properties of Nanotubes. J. Chem. Inf. Model. **2013**, 53, 773-782.

(69) Wang, X. Z.; Yang, Y.; Li, R. F.; McGuinnes, C.; Adamson, J.; Megson, I. L.; Donaldson, K.: Principal Component and Causal Analysis of Structural and Acute in vitro Toxicity Data for Nanoparticles. Nanotoxicology **2013**, 1-31.

(70) Quintero, F. A.; Patel, S. J.; Muñoz, F.; Sam Mannan, M.: Review of Existing QSAR/QSPR Models Developed for Properties Used in Hazardous Chemicals Classification System. Ind. Eng. Chem. Res. **2012**, 51, 16101-16115.

(71) Lubinski, L.; Urbaszek, P.; Gajewicz, A.; Cronin, M.; Enoch, S.; Madden, J.; 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.

(72) Richarz, A. N.; Cronin, M.; Madden, J.; Lubinski, L.; Mokshina, E.; Urbaszek, P.; Puzyn, T.; Kuz'min, V.: Toxicity of nanomaterials: Availability and suitability of data for the development of in silico models. Toxicol. Lett. **2013**, 221, S246-S246.

(73) Gütlein, M.; Helma, C.; Karwath, A.; Kramer, S.: A Large - Scale Empirical Evaluation of Cross - Validation and External Test Set Validation in (Q) SAR. Molecular Informatics **2013**, 32, 516-528.

(74) Chirico, N.; Gramatica, P.: Real external predictivity of QSAR models: how to evaluate it? Comparison of different validation criteria and proposal of using the concordance correlation coefficient. J. Chem. Inf. Model. **2011**, 51, 2320-2335.

(75) Konovalov, D. A.; Llewellyn, L. E.; Vander Heyden, Y.; Coomans, D.: Robust cross-validation of linear regression QSAR models. J. Chem. Inf. Model. **2008**, 48, 2081-2094.

(76) Golbraikh, A.; Tropsha, A.: Beware of q^2! Journal of Molecular Graphics and Modelling **2002**, 20, 269-276.

(77) Golbraikh, A.; Shen, M.; Xiao, Z.; Xiao, Y.-D.; Lee, K.-H.; Tropsha, A.: Rational selection of training and test sets for the development of validated QSAR models. J. Comput.-Aided Mol. Des. **2003**, 17, 241-253.

(78) Van Leeuwen, C. J.; Patlewicz, G. Y.; Worth, A. P.: Intelligent testing strategies. In Risk Assessment of Chemicals; Springer, 2007; pp 467-509.

(79) Frank, R. A.; Sanderson, H.; Kavanagh, R.; Burnison, B. K.; Headley, J. V.; Solomon, K. R.: Use of a (quantitative) structure–activity relationship [(Q) Sar] model to predict the toxicity of naphthenic acids. J. Toxicol. Environ. Health, Part A **2009**, 73, 319-329.

(80) Benigni, R.; Netzeva, T. I.; Benfenati, E.; Bossa, C.; Franke, R.; Helma, C.; Hulzebos, E.; Marchant, C.; Richard, A.; Woo, Y.-T.: The expanding role of predictive toxicology: an update on the (Q) SAR models for mutagens and carcinogens. Journal of Environmental Science and Health Part C **2007**, 25, 53-97.

(81) Patlewicz, G.; Chen, M.; Bellin, C.: Non-testing approaches under REACH–help or hindrance? Perspectives from a practitioner within industry. SAR QSAR Environ. Res. **2011**, 22, 67-88.

(82) Jaworska, J. S.; Comber, M.; Auer, C.; Van Leeuwen, C.: Summary of a workshop on regulatory acceptance of (Q) SARs for human health and environmental endpoints. Environ. Health Perspect. **2003**, 111, 1358.

(83) Puzyn, T.; Gajewicz, A.; Leszczynska, D.; Leszczynski, J.: Nanomaterials–the Next Great Challenge for Qsar Modelers. In Recent Advances in QSAR Studies; Springer, 2010; pp 383-409.