

## Computational Models in Chemical Safety Assessment

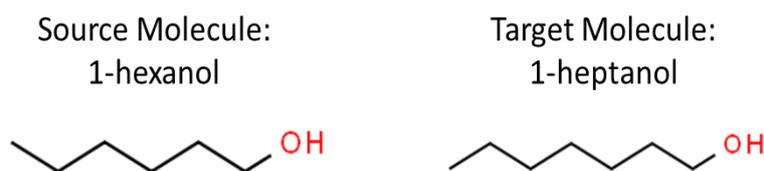
How are computational models used in chemical safety assessment? This is a vital, but complex, question for 21<sup>st</sup> Century toxicology: how can they be used to determine whether a particular use of a chemical is safe? To understand how toxicology uses computational models it is important to appreciate the current needs of safety assessment, and the vision of how computational methods, may ultimately be used to replace the animal tests, whilst providing the same or an improved level of protection for humans and the environment. Here we explore the concept and application of computational models, as well as their possibilities for the future.

### What is computational modelling of toxicity?

Chemical safety assessment has traditionally relied on experimental data generated predominantly from animal testing. The results from such tests are used to identify any potential hazard and, where possible to assess potency, so that informed decisions about safety can be made. There is a growing use of modelling and other computational techniques to supplement and even replace, the use of animals to provide information for safety assessment. These models, often called computational or “*in silico*” approaches, utilise knowledge gained from existing test data to help identify the probable effects of new molecules for which data are lacking. There is a broad spectrum of computational approaches. On the whole, all models have one thing in common: they identify the aspects chemical structure responsible for an effect and apply that knowledge to predict effects for substances with no data. This is comparable to a meteorologist taking measurements e.g. air and sea temperature, atmospheric conditions etc., to make a model of the weather and then making the “weather forecast” on the basis of this information.

The computational models which are applied in chemical safety are routinely used to predict toxicity and pharmacokinetic effects. One fundamental premise of these models is that similar molecules will have similar effects. The technique of “*read-across*” uses this approach to predict activity for structurally similar molecules i.e. reading across the activity from a molecule with known activity to one with few or no data. A simple example is shown in Figure 1. Read-across approaches commonly use ‘*structure-activity relationships*’ (SARs), where molecular fragments are associated with a particular activity. These SARs can form the basis of structural rules, which are coded computationally into ‘*structural alerts*’ which help to identify molecular features associated with hazardous properties. Quantitative models are also commonly used, these relate aspects of chemical structure captured by physico-chemical properties and molecular descriptors to existing toxicity or effect data; these are termed quantitative structure-activity relationships (QSARs). As larger and more varied sets of chemicals have been compiled, there has also been a growth in the application of machine learning (ML) through artificial intelligence (AI) to gain an understanding of toxicity and to model it.

## Public Statement. [British Toxicology Society](#)

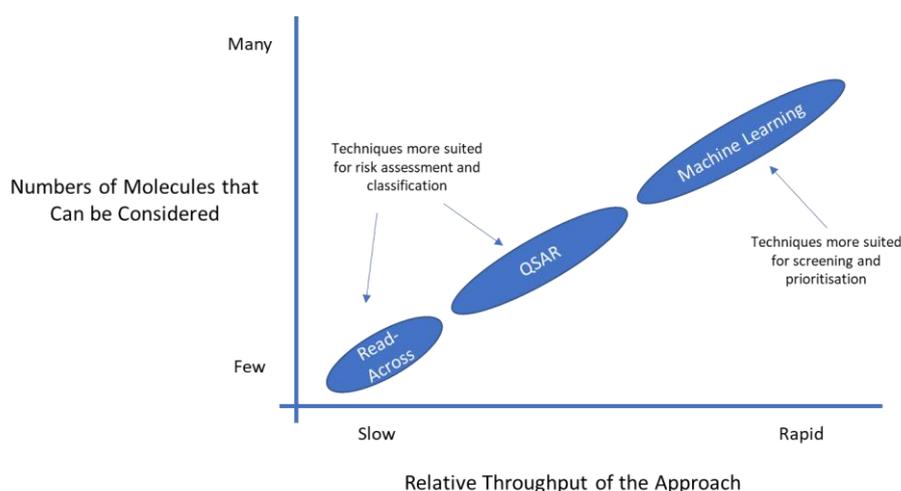


**Figure 1.** An example of read-across to fill data for 1-heptanol using data for 1-hexanol. This one-to-one read-across is termed an analogue approach. Toxicity of a 'data poor' (target) molecule is informed by the "data rich" (source) molecule. The molecules are similar in terms of the chemical functional group (aliphatic alcohol OH) and differ by only one carbon atom. The possible effects of the differences in chemical structure, e.g. on aqueous solubility, and the possible effect on toxicity should be considered

### How and where are computational models of toxicity used?

As with the toxicological information they intend to supplement or replace, predictions from computational models have a wide number of uses in chemical safety assessment. Almost all industrial sectors apply computational models to assess safety, most notably in the personal care products, pharmaceutical, biocide and chemical industries. Such modelling is seen as a fundamental method to reduce costs and improve the efficiency of product development as well as one of several tools to replace animal testing, as for instance in the cosmetics sector. The specific uses of models range from the rapid screening of large libraries of compounds in the early stage of product development (e.g. to remove potentially toxic molecules), through providing information for hazard assessment on single compounds, to providing safety information which replaces an animal test.

It's an important and often overlooked fact that the use and expectations of a prediction of toxicity will dictate type of the model that is used. ML predictions are likely to be highly useful for eliminating potential toxic molecules when screening large chemical inventories, whereas read-across has found favour where greater certainty is required for complex effects. The relative throughput (speed) of the different computational toxicology techniques is shown in Figure 2. To better understand the role of different models, we need to consider how to determine whether a prediction is reliable.

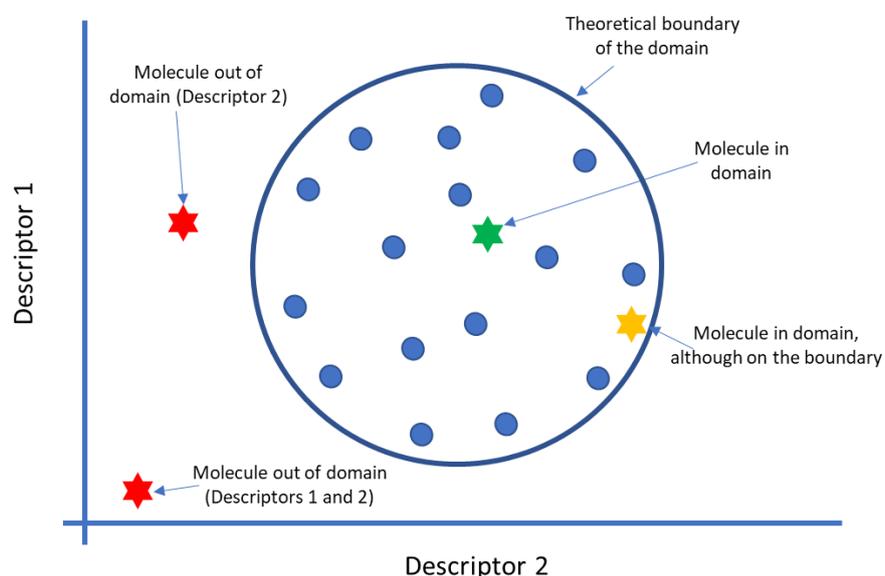


**Figure 2.** The relative throughput of computational techniques dictates the numbers of molecules that can be assessed and influences how they can be used. Read-across is often relatively slow due to the requirement of data gathering and documentation. Machine learning algorithms can be run rapidly over large chemical inventories, but with potentially great uncertainty in the output.

### Are predictions of computational models of toxicity reliable?

The reliability, or accuracy, of a model's prediction is key when making a decision on chemical safety based on it. For a scientific discipline based on making decisions using experimental evidence, toxicology requires the same confidence from predictions that would be provided by experimental evidence (if not more!!). In order to use a prediction of toxicity from a computational model, we need to determine if it is "acceptable". To understand how we determine "acceptability" of a prediction, we need to separate out two aspects. The first is an assessment of the model itself in terms of its reliability, performance and suitability for purpose. The second is whether the model is appropriate to make a prediction. For instance, we may be able to develop a high-quality model for an effect brought about by a particular class of pharmaceuticals. However, regardless of how good the model is in the initial application, it may not be appropriate to use the model for other chemical or pharmacological classes.

The most stringent use of predictions is probably to replace an animal test, for instance for regulatory purposes. To determine whether a prediction is acceptable, clear criteria are applied to determine the "validity" of the model as well as to assess whether it is appropriate to make a prediction for the chemical in question e.g. whether the compound is within the "applicability domain" of the model. Figure 3 illustrates how molecules may be in or out of the applicability domain of a model. For all uses of models to predict toxicological effects, there is an expectation that the model should be characterised in terms of the chemicals, activity data and descriptors within the model, the algorithm or hypothesis of the model or read-across, and the general performance of the approach. For some purposes, such as regulatory use e.g. within [REACH](#), the level of justification of a model and prediction is expected to be high, due to the legislative requirements underpinning the use of *in silico* predictions as an adaptation of a regulatory test.



**Figure 3.** Plot of the descriptors for the molecules in a theoretical QSAR – the training set (on which the model is based) is shown in blue circles. Test set molecules are shown as stars either within the domain of the model (green), on its boundary and hence possibly more uncertain (yellow) or out of the domain (red). A prediction should only be made for molecules that can be shown to be within the domain of the model.

**If computational models of toxicity are reliable, why are chemicals still being tested on animals?**

There are many reliable computational models of toxicity and growing expertise in how to apply them. However, the concept of acceptability has been one of the main stumbling blocks for the greater uptake of computational models in chemical risk assessment, particularly for regulatory use. For read-across and QSAR there is much guidance, but so far few concrete examples and case studies demonstrating how, and when, *in silico* approaches may be acceptable. In addition, the current validation principles, especially for QSARs, are robust but do not yet capture the possibilities of using some newer computational methods. Thus, whilst there are potential usable models, animal tests may still be required to meet regulatory and scientific requirements, until the model developer and user can demonstrate that predictions are appropriate, reliable and fit for purpose.

**Where next for computational modelling of toxicity?**

The vision for computational models of toxicity is potentially endless, with increased use and uptake anticipated. 21<sup>st</sup> Century toxicology puts computational modelling at the heart of chemical safety assessment both as a standalone tool and part of a suite of approaches that can be used to provide information. Looking further ahead, there will undoubtedly be a need to more effectively model big data. AI will be used in the prediction of toxicity, especially when supported by appropriate big data – but its implementation requires a sensible and rational approach, so that the expectations (the ultimate replacement of animal tests) and requirements (in terms of regulatory acceptance) of those who use the predictions can be met. Much can be learnt from experience: the QSAR community have used what are now termed ML approaches regularly, since computational power allowed in the 1980s; and neural networks since the early 1990s. ML approaches are routinely applied to problems where rapid assessment of large numbers of chemicals is required, often with little assessment or evaluation of individual predictions e.g. screening out molecules with potential for a particular type of toxicity, from compound libraries.

The biggest challenge facing computational toxicology is clearly the ambition to replace animal tests. This is currently possible for some chemicals and some endpoints, such as mutagenicity and skin sensitisation – acceptance for regulatory use has demonstrated this. However, much remains to be done in defining, accepting, and addressing, the limitations of our computational methods. Yes, chemical structure-based predictions can replace animal tests in some situations, but in many instances, they are currently most likely to be used as part of a weight of evidence which includes mechanistic, *in vitro* and other types of information. Looking to the future, knowledge of chemistry will be harnessed through computational approaches and models to simulate and predict toxic effects of chemicals, which will enable more rapid, accurate, ethical and cheaper safety assessment.