Approaches for addressing data needs in risk assessment

Fact sheet series: Topics in risk assessment of substances under the *Canadian Environmental Protection Act, 1999* (CEPA 1999)

A common issue in risk assessment is a lack of available experimental data on the characteristics of a substance (for instance, the substance's physical-chemical or toxicological properties, or its environmental fate). In the absence of available experimental data, the approaches described below can be relied upon for risk assessment purposes. More information on the data collection process can be found in the information gathering fact sheet.

Experimental data generation

In some cases, it may be possible to conduct studies to generate the experimental data needed for an assessment.

Stakeholders may help to reduce the uncertainty in an assessment by sharing information on a substance that is not publicly available, or by conducting testing on a substance. The government can also require the generation of specific types of experimental data. CEPA 1999, under section 71(1)(c), enables the Minister of the Environment to require persons involved in the manufacture or import of a substance that is suspected to be toxic to conduct testing and submit the results. Before a new substance is introduced into the Canadian market, section 84(1)(c) of CEPA 1999 enables the Minister of the Environment to request the results of any testing that is considered necessary for

assessing a substance that is toxic or capable of becoming toxic. However, a company may decline the request for testing if they decide not to manufacture or import the substance.

Experimental data may also be generated by researchers at Environment and Climate Change Canada and Health Canada. Both departments have programs that conduct research to support the assessment and management of substances.

Predictive tools

It is not always possible to generate new data when needs are identified for an assessment. There is also an ongoing effort to replace, reduce and refine the use of animal toxicity testing and to advance alternative scientific approaches. The use of predictive tools to fill data needs is an effective alternative option.

Predictive tools used in the risk assessment of substances are generally based on the assumption that substances that are structurally similar will have similar physical-chemical properties, behave similarly in humans, other organisms, and in the environment, and elicit similar toxic effects. Therefore, experimental data for similar substances can be used to predict the behaviour of a substance that lacks data. This can be done in several ways.

Use of analogues: read-across and category approaches

The use of analogues and read-across in risk assessment fact sheet provides a detailed description of the read-across and category approaches. A read-across approach considers experimental analogue data to address data needs for a target substance. Data for one or more analogue substances with similar

chemical structure or properties may be used to estimate missing information for the target substance being assessed. While similar to the read-across approach, a category approach can be used to address data needs within a group of substances that are similar (a category).

Computational modelling, including (Quantitative) Structure-Activity Relationships [(Q)SARs]

Computational modelling can be used to predict key data endpoints that are essential for carrying out a risk assessment. For example, (Q)SAR is a method that establishes a relationship between structural characteristics or properties of a substance and its behaviour. Multiple substances are used as what is called a training set to identify and establish relationships between certain descriptors or properties of the substances and their toxicity/activity. Algorithms (such as statistical methods) are then used to build (Q)SAR models that make predictions on the toxicity/activity of substances outside of the training set. Models may base their predictions on:

- relationships seen between physical-chemical properties and behaviour of the substances, which are often based on a simple regression equation (for example, an increase in toxicity with an increased octanol/water partition coefficient)
- relationships seen between various structural fragments of a substance and its properties or behaviours (for example, with each additional carbon atom in an alkane, the boiling point can be expected to increase by a certain amount)
- how certain types of substances behave, where rules and exceptions are built into a model to predict the activity of a substance (for example, if a substance has a functional

group known to bind to DNA and does not have other characteristics that would cause it to detoxify, then it would be predicted as a potential mutagen)

The use of predictive models to generate data can be quick and inexpensive, but can also have its limitations. Models may not be able to make accurate predictions for substances that are outside the “domain of applicability”. For instance, when the training sets (which are chemicals used to create the model) do not include substances similar to the substance of interest. Certain types of substances can be particularly “model-difficult” for certain parameters in that the model does not easily establish relationships between structural characteristics or properties of a substance and its behaviour. Some examples of difficult to model substances include polymers, surfactants, ionizable substances, pigment and dye structures, substances with high octanol-water partition coefficients, and unknown or variable composition complex reaction products and biological materials (UVCBs). Confidence in the model prediction for a particular substance is always taken into consideration when deciding whether to use modelled data in a risk assessment, and how much weight to give it in making a decision.

In a risk assessment, it is possible to make use of any combination of these approaches to address data needs. Where data are available through more than one approach, these different lines of evidence are considered in a weight of evidence approach.