10 Predicting Carcinogenicity Potential Using OncoLogic™

Definition of Cancer
Cancer is defined as the uncontrolled dividing and growth of cells. Uncontrolled cell division and growth can be caused by mutations, increase in cell proliferation, decrease in cell death, or loss of homeostatic control. The two general mechanisms by which a chemical can induce cancer include:

- **Genotoxic** which is the default mechanism in which the chemical interacts with the cell’s DNA to cause mutation(s) in genes; and
- **Non-genotoxic** by a variety of mechanisms.

Three Stages of Carcinogenesis
Carcinogenesis (the production or development of cancer) is a multistage/multistep process. The three stages or steps in the process, shown in the figure at the right are:

- **Initiation**: Mutation converts normal cells to preneoplastic cells
- **Promotion**: Expansion of preneoplastic cells to benign tumors
- **Progression**: Transformation of benign to invasive malignant tumors

**OncoLogic™ concern levels**
- L = Low
- M = Marginal
- LM = Low – Moderate
- MH = Moderate – High
- H = High
A potent carcinogen is one that acts directly on all three stages, as shown in the center of the diagram, and has a high cancer concern level. A weak carcinogen acts directly on one stage and indirectly on other stages, as shown in the diagram where only two stages overlap.

**Performing a Cancer Screen Under the Sustainable Futures Initiative**

This chapter describes two methods for performing a cancer screen under the Sustainable Futures initiative:

- Method 1. Use experimental data on the chemical or an analog
- Method 2. Use computer-based expert system (OncoLogic™) to predict carcinogenicity

**10.1 Perform Cancer Screen Using Experimental Data on Chemical or Analog**

Generally there are two types of experimental cancer studies: (1) laboratory studies and (2) epidemiological studies.

**10.1.1 Laboratory Studies** are conducted in controlled environments using laboratory animals. Animals dosed with test substance or with vehicle (e.g., water or corn oil) for majority of their life. Tissues are examined for tumors at the end of the exposure period (or in animals that die prior to the scheduled sacrifice). The number of animals in treatment groups with tumors is compared with the number of animals in control group(s) with tumors in same tissue.

**10.1.2 Epidemiological Studies** are based on review of human populations. The incidences of cancer in human populations that have been accidentally or inadvertently exposed to a substance are compared with incidences of cancer unexposed populations. Interpreting results is often complicated due to lifestyle and genetic differences of the population groups.

**10.1.3 Interpreting Experimental Data**

*Positive Studies*

Indications of positive studies are a statistically significant increase in number of animals with cancer at one or more dose(s); statistically significant trend in number of animals with tumors; and the presence of rare tumor(s). Note that the default assumption is that instances of tumors seen in experimental animals are relevant to human health, however some exceptions exist. For example some kidney tumors in male rats are caused by protein that is not found in human kidneys at appreciable concentrations.

*Negative Studies*

Evaluate the study adequacy using this checklist:

1. Were a sufficient number of animals dosed for a sufficient length of time?
2. Were animals given appropriate doses?
3. Were enough tissues examined microscopically?
4. Is the exposure route relevant (consider absorption rates for oral, inhalation, dermal exposure)

Additional guidance for evaluating study adequacy can be found at [http://www.epa.gov/docs/OPPTS_Harmonized/870_Health_Effects_Test_Guidelines/Series/](http://www.epa.gov/docs/OPPTS_Harmonized/870_Health_Effects_Test_Guidelines/Series/)
10. Predicting Carcinogenicity Potential Using OncoLogic™

10.1.4 Experimental Cancer Data Can Be Found In These Sources

Here is a list of the major online sources of experimental cancer data. Additional data sources are listed in Appendix C.
- National Toxicology Program (NTP) http://ntp-server.niehs.nih.gov/
- Environmental Health Criteria (EHC) http://www.inchem.org/
- Chemical Carcinogenesis Research Information System (CCRIS) searchable through TOXNET http://toxnet.nlm.nih.gov/
- Gold database http://potency.berkeley.edu
- ACToR – Aggregated Computational Toxicology Resource (EPA) http://www.epa.gov/actor/ or http://www.epa.gov/comptox/

10.1.5 Search for Experimental Cancer Data on Isodecyl Acrylate


10.2 Use OncoLogic™ to Predict Cancer Concern

10.2.1 What is OncoLogic™?

OncoLogic™ is a Windows®-based computer program that evaluates the likelihood that a chemical may cause cancer. OncoLogic™ has been peer reviewed and the subject of external validation.

10.2.3 Downloading OncoLogic™

OncoLogic™ Version 7.0 can be downloaded from the Sustainable Futures web site at http://www.epa.gov/oppt/sf/pubs/oncologic.htm. Additional documentation on OncoLogic™ is available on the OncoLogic™ download page including:
- **OncoLogic™ Quick Start Tutorial** which includes step-by-step instructions on running OncoLogic™
- **OncoLogic™ User Manual** which offers tips on running the 48 specific organic subsystems.
- **Frequently Asked Questions** on OncoLogic™
- **Running the model** has summary information on OncoLogic™ and flow diagrams showing each module.
- **Interpreting results** includes pages from the Interpretive Guidance Document on interpreting OncoLogic™ results; and
- **OncoLogic™ Cancer Assessment Presentation** used during Sustainable Futures hands-on training.

10.2.4 Challenges in Predicting Carcinogenicity

Developing Quantitative Structure-Activity Relationship (QSAR) methods to predict carcinogenicity is difficult for many reasons. The QSAR would have to be highly complex and dependent on the mechanisms of action. The data needed to develop QSARs are scarce and have a high degree of variability. The necessary steps involving feedback and validation would be problematic, and the QSAR would require an integrative approach.

The SAR methods within OncoLogic™ differ from those in EPISuite and ECOSAR because it is difficult to relate specific physical / chemical properties to carcinogenicity. Many properties can have multiple
possible effects on carcinogenicity. As discussed previously there are multiple stages in the expression of carcinogenicity. Often a necessary step is metabolism from a parent compound that is not a carcinogen to a carcinogenic intermediate. Isomers that have very similar properties may have dramatically different cancer concerns (as shown in the image below). As a result there are no all-encompassing descriptors have been identified for carcinogenicity even within many chemical classes.

Carcinogenicity of a chemical may be drastically different for chemicals with similar chemical/physical properties. For example a change in the location of the amine group in beta-naphthylamine to form alpha-naphthylamine (shown to below) and the carcinogenic potential drops from high to marginal.

The critical factors to consider when developing QSARs for carcinogenicity include:

Electronic and Steric Factors
- Resonance stabilization
- Steric hindrance
- Molecular size and shape

Metabolic Factors
- Blocking of detoxification
- Enhancement of activation

Mechanistic Factors
- Electrophilic vs. receptor-mediated
- Multistage process

Physical / chemical Factors
- Molecular weight
- Physical state
- Solubility
- Chemical reactivity

10.2.5 How Does OncoLogic™ Work?
OncoLogic™ is an expert system that mimics the thinking and reasoning of human experts using knowledge based rules for chemical classes to predict cancer concern. It assigns a baseline concern level ranging from low to high (shown below) by evaluating how substituents on the chemical may affect carcinogenicity.

<table>
<thead>
<tr>
<th>OncoLogic Concern</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low (L)</td>
<td>Unlikely to be carcinogenic</td>
</tr>
<tr>
<td>Marginal (mar)</td>
<td>Likely to have equivocal carcinogenic activity</td>
</tr>
<tr>
<td>Low – Moderate (LM)</td>
<td>Likely to be weakly carcinogenic</td>
</tr>
<tr>
<td>Moderate (M)</td>
<td>Likely to be a moderately active carcinogen</td>
</tr>
<tr>
<td>Moderate – High (MH)</td>
<td>Highly likely to be a moderately active carcinogen</td>
</tr>
<tr>
<td>High (H)</td>
<td>Highly likely to be a potent carcinogen</td>
</tr>
</tbody>
</table>
Benefits of OncoLogic™

OncoLogic™ has numerous benefits including:

- Allows non-experts to reach scientifically supportable conclusions that are backed by a mechanistic understanding/rationale and capable of generating testable hypotheses,
- Allows the sharing of knowledge by incorporating decades of cancer research and the practical experience of a team of domain experts, in predicting the carcinogenic potential of chemicals by SAR analysis,
- It is a flexible system capable of incorporating both chemical and biological information to predict potential carcinogenicity of a variety of chemicals.
- Expedites the decision making process, reducing errors and inconsistencies, and
- Formalizes knowledge rules for cancer hazard identification.

Limitations

OncoLogic™ does have certain limitations that users should understand, specifically:

- It is primarily designed to predict potential carcinogenicity of industrial chemicals and therefore may not work as well for pharmaceuticals;
- No batch-mode function exists for entering several chemicals at one time;
- Chemical cannot be entered using structural formats like SMILES; and
- In order to use OncoLogic™ correctly, the user must have a basic knowledge of organic chemistry and ability to place chemicals in the appropriate chemical class.

What do I need to know to run OncoLogic™?

A user should be familiar with organic chemistry and know certain characteristics of the chemical of interest including: structure (including subunits present), physical/chemical properties (stability, etc.), biological and mechanistic information, as well as possible routes of exposure. For example, there are specific inputs required for fibers, metals and polymers, and these are explained in greater detail in the OncoLogic™ Users Manual. The general public is advised to seek assistance from chemists, carcinogenesis experts, and other environmental health specialists when using OncoLogic™.

10.2.6 Design Aromatic Amine Dyes with Lower Carcinogenic Potential Using Knowledge Rules

Start with benzidine (left) which has a high cancer concern and using Knowledge Rules you can make structural changes that will lower the cancer concern. Five examples are provided below along with the changes made to the original molecule and the effect on the cancer concern level.

Example 1: Introduce bulky substituent(s) ortho to amino / amine-generating group(s). Result provides steric hindrance to inhibit bioactivation.
Concern = Marginal
10. Predicting Carcinogenicity Potential Using OncoLogic™

Example 2: Introduce bulky N-substituent(s) to amino / amine-generating group(s). Result makes it a poor substrate for the bioactivation enzymes. Concern = Marginal

Example 3: Introduce bulky groups ortho to intercyclic linkages. Result distorts the planarity of the molecule making it a poor substrate for the bioactivation enzymes. Concern = Marginal

Example 4: Replace electron-conducting intercyclic linkages by electron-insulating intercyclic linkages. Two results: (1) Reduce length of conjugation path and thus the force of conjugation, which facilitates departure of acyloxy anion; and (2) Less resonance stabilization of electrophilic nitrenium ion. Concern = Marginal

Example 5: Substitution of ring with hydrophilic and/or electron withdrawing groups. Two results: (1) Render molecule more water-soluble thus reducing absorption and accelerating excretion; and (2) Makes amines less nucleophilic. Concern Level = Low

10.3 Running OncoLogic™ Modules

OncoLogic™ evaluates potential carcinogenicity by using mechanism-based structure-activity relationship (SAR) analysis. The chemical of interested is compared to structurally related compounds with known carcinogenic activity. Structural moieties or fragments that may contribute to carcinogenic activity through a perceived or postulated mechanism are identified and evaluated. The structural analysis arm consists of four modules (described below), including the Organics module, Metals module, Polymers module, and Fibers module.

10.3.1 Organics Module (right)
- Enter information on chemical identity
- Choose appropriate chemical class
- Enter chemical name, CAS#, or chemical structure
- Select chemical class
- 48 total
- Description in Manual
- Select “Help” to view sample structures
- Absence of structure in OncoLogic provides suggestive, but not definitive, evidence of low cancer concern
10.3.2 Metals Module (below)
- Similar to running the organics module
- Pick the metal to be evaluated
  - OncoLogic™ will then either ask a series of questions needed to evaluate the chemical or provide a database of related compounds

Information Needed to Run the Metals Module
- Nature/form of the metal / metalloid
  - Organometal, metal powder
- Type of chemical bonding (e.g., organic, ionic)
- Dissociability / solubility
  - Valence / oxidation state
- Crystalline or amorphous
- Exposure scenario
- Breakdown products (e.g., organic moieties)

10.3.3 Polymers Module
- Polymer must consist of covalently linked repeating units and have a number average molecular weight >1000
- OncoLogic™ asks a series of questions designed to aid in evaluation of carcinogenicity of the polymer

Information Needed to Run the Polymers Module
- Percentage of polymer with molecular weight <500 and <1000
- Percent of residual monomer
- Identification of Reactive Functional Group(s)
- Solubility
- Special features
  - Polysulfation, "water-swellability"
- Exposure route
- Breakdown products (e.g., hydrolysis)

10.3.4 Fibers Module
- Evaluations are based on physical dimensions and physicochemical properties
  - Physical dimensions
    - Diameter, length, aspect ratio
  - Physicochemical properties
    - High density charge, flexibility, durability, biodegradability, smooth and defect-free surface, longitudinal splitting potential
    - Presence of high MW polymer, low MW organic moiety, metals/metalloids

Information Needed to Run the Fibers Module
- Relevant manufacturing / processing / use information
- Crystallization, thermal extrusion, naturally occurring, unknown method
10.4 Functional Arm Analysis

Functional analysis integrates available mechanistic / non-cancer studies on the chemical in order to predict the potential for the chemical to be a tumor initiator, promoter, or progressor. The image at the right shows how the Functional Arm scores and integrates information to identify the potential for the chemical to be an initiator, promoter, or progressor.

Results from the functional analysis can be used to provide support for the results of the structural analysis, or can be used as an independent method of analysis. Functional analysis can be used to identify data gaps and relevancy of animal data to human. The scoring and weighting factors may be subject to changes as new knowledge becomes available. The structural and functional analyses must be performed separately.

The types of data that are included in the Functional Arm Analysis include:
- Oncogene/Tumor suppressor gene data,
- Transgenic rodent test data,
- Genotoxicity and DNA reactivity data,
- Epigenetic toxicity data,
- Subchronic toxicity data, and
- Additional mechanistic information, testing conditions, and dosages.

The Integrated Report produced by the Functional Arm Analysis includes:
- Summary of test data entered,
- Integrated score,
- Concern level,
- Test interpretation, and
- Data gap and recommendations.
10.5 Interpreting Results from OncoLogic™

The table below summarizes how results from OncoLogic™ can be interpreted in a Sustainable Futures Summary Assessment.

<table>
<thead>
<tr>
<th>OncoLogic™ Concern Level</th>
<th>Sustainable Futures Concern Level</th>
<th>Definition</th>
<th>Proceed to Risk Screen?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>Low</td>
<td>Unlikely to be carcinogenic</td>
<td>No</td>
</tr>
<tr>
<td>Marginal</td>
<td>Further Research Needed</td>
<td>Likely to have equivocal carcinogenic activity</td>
<td>Additional information is needed</td>
</tr>
<tr>
<td>Low – Moderate</td>
<td>Moderate</td>
<td>Likely to be weakly carcinogenic</td>
<td>Yes</td>
</tr>
<tr>
<td>Moderate</td>
<td>Moderate</td>
<td>Likely to be a moderately active carcinogen</td>
<td>Yes</td>
</tr>
<tr>
<td>Moderate – High</td>
<td>High</td>
<td>Highly likely to be a moderately active carcinogen</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td></td>
<td>Highly likely to be a potent carcinogen</td>
<td>Yes</td>
</tr>
</tbody>
</table>

10.6 Running the Sample Chemical Isodecyl acrylate in OncoLogic™

Here are the steps in evaluating the sample chemical Isodecyl acrylate in OncoLogic™. Screen captures are not included here however similar images are available in the OncoLogic™ Quick Start Tutorial and OncoLogic™ User Manual.

1. Start OncoLogic™
2. Click on “New SAR Analysis”, then “Organics”
3. When the wizard pops up click on “New Compound”
4. Enter an identifying Code “1330616” and Description “Isodecyl acrylate” of the substance
5. Select “Acrylates” as the Chemical Class.
6. The wizard will ask you if the molecule to which the reactive functional group is attached is known to be NOT bioavailable by the anticipated route(s) of exposure, and you select “NO”
7. Next enter the Molecular Weight of Isodecyl acrylate which is 212
8. Enter the expected exposure routes from the 4 choices - Inhalation, Injection, Oral, Dermal - and you select Inhalation and Dermal
9. Next you select the acrylate skeleton to build the molecule
10. Use the drawing tools to build the molecule
11. Click on Done
12. You will be asked if there is another reactive functional group attached to the compound – select NO.
13. The final Justification Report is presented and you can print or save it as RTF (rich text format)
Important Notes on Selecting Class (Step 5 above)

- Selecting the appropriate class for your chemical can be the most difficult step in running OncoLogic™. If, after you select a class, you click on the "Help" you will get information on that class, including structures within that class (shown below). This information can help you determine if your chemical does belong in that class.
- If the chemical you are evaluating in OncoLogic™ belongs in two or more classes, evaluate each class and use worst case to evaluate the molecule.
- If the chemical does not fit well within any of the classes in OncoLogic™ it is probably not within the domain of the model and should not be evaluated using OncoLogic™.
10. Predicting Carcinogenicity Potential Using OncoLogic™

10.6.1 OncoLogic™ Justification Report
The Justification Report for Isodecyl acrylate is shown below. OncoLogic™ predicts a Marginal cancer concern for the sample chemical.

![OncoLogic Justification Report]

**SUMMARY:**
- **CODE NUMBER:** 1310616
- **SUBSTANCE ID:** Isodecyl acrylate

The final level of carcinogenicity concern for this acrylate when the anticipated route of exposure is inhalation or injection is MARGINAL.

**JUSTIFICATION:**
An acrylate is a potential alkylating agent which may bind, via Michael addition, to key macromolecules to initiate/exert carcinogenic action. The alkylating activity of acrylates can be substantially inhibited by substitution at the double bond, particularly by bulky or hydrophobic groups. The nature and molecular size/shape of the molecule to which the acrylate is attached may also play a role in affecting the overall activity of the compound.

The acrylate is stable and has a baseline level of concern of LOW-MODERATE.

The molecule to which the acrylate is attached, which is denoted as R1, is expected to reduce the level of concern.

Therefore, the level of concern is reduced to LOW.

In general, inhalation and injection provide the best chance of delivering the largest possible amount of direct-acting reactive chemicals to target tissue because of a lesser absorption barrier and better chance of avoiding denaturation by protease nucleophiles such as glutathione. Since the exposure to this compound is expected to be by only one or both of these routes, the level of concern is expected to be increased.

Therefore, the level of concern is raised to MARGINAL.

The final level of concern when the anticipated route of exposure is inhalation or injection is MARGINAL.

10.6.2 Entering Results into the Sustainable Futures Worksheet
You will enter the OncoLogic™ results into the Sustainable Futures Summary Assessment Worksheet on page 4:
- **Experimental data** = Low by analogy to isoctyl acrylate (Gordon el a1 1991)
- **OncoLogic Results** = Marginal
- **Overall Hazard Concern for Carcinogenicity** = Low

<table>
<thead>
<tr>
<th>CANCER HEALTH EFFECTS:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experimental data</strong></td>
</tr>
<tr>
<td><strong>OncoLogic Results</strong></td>
</tr>
<tr>
<td><strong>Overall Hazard Concern for Carcinogenicity</strong></td>
</tr>
</tbody>
</table>
10.7 Validation and Evaluations of OncoLogic™

The details of the validation of OncoLogic™ are discussed in Woo YT, et al 1995; and Woo YT and DY Lai 2005 (full citations are at the end of this chapter). Essentially, beyond internal validation and crosschecking, external peer reviewing and prospective validation were conducted. OncoLogic™ was peer reviewed at the developmental stage by external domain experts and, after completion of versions 2.0 and 4.0, by two international peer review panels of domain experts. In addition, the scientific bases of rule packages for a number of classes of chemicals were published in peer-reviewed open literature (e.g., Lai DY, et al 1996A and 1996b; Woo YT and Lai DY, 2001).

The OncoLogic™ team participated in two international, prospective predictive exercises sponsored by NTP/NIEHS to evaluate the capabilities of various methods to predict the outcome of cancer bioassays several years before the studies were completed. In the first exercise, focusing on 1 of 8 aromatic amines, OncoLogic™ achieved a high degree of accuracy, as shown in the table below (Woo YT, et al 1995).

### OncoLogic™ Prediction vs. NTP Bioassays for Aromatic Amines and Related Compounds

<table>
<thead>
<tr>
<th>NTP #</th>
<th>Chemical</th>
<th>Bioassay Results</th>
<th>Oncologic™ Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>4,4′-Diamino-2,2′-stilbene disulfonic acid</td>
<td>N/N N/N --</td>
<td>L</td>
</tr>
<tr>
<td>42</td>
<td>p-Nitroaniline</td>
<td>NT E/N Eq mar</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>p-Nitrobenzoic acid</td>
<td>N/S N/N + mar</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>p-Nitrophenol</td>
<td>NT N/N --</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>4-Hydroxyacetanilide</td>
<td>N/E N/N Eq LM</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>2,4-Diaminophenol dihydrochloride</td>
<td>N/N S/N + M</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>3,3′-Dimethylbenzidine</td>
<td>C/C NT + HM</td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>o-Nitroanisole</td>
<td>C/C C/C + HM</td>
<td></td>
</tr>
</tbody>
</table>

Results Codes

- **C** = Clear evidence of carcinogenicity
- **S** = Some evidence of carcinogenicity
- **N** = No evidence of carcinogenicity
- **NT** = Not tested
- **+** = At least one test = C or S
- **Eq** = No C or S, and E must appear at least once
- **--** = No C, S, or E

In a second exercise on 30 chemicals of diverse structure, the OncoLogic™ approach (Woo YT et al 1997) was rated as one of the best performers (Benigni R, Zito R. 2004). A recent external validation by U.S. FDA (Mayer J, Cheeseman MA, Twaroski ML 2008) showed that, within the limitations of the method, the predictive accuracy of OncoLogic™ exceeded 90% for the batch of chemicals the Agency was interested in. It should be cautioned, however, that the predictive accuracy of OncoLogic™ is expected to vary from batch to batch depending upon the structural diversity of chemicals relative to the underlying knowledge basis of OncoLogic™.

10.8 Major References on OncoLogic™


Intentionally Left Blank