Curating Chemical Use Categories and Exposure Predictions to Inform Chemical Assessment

Victoria Hull, M.S.
Inotiv, Inc., contractor supporting the NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)

ASCCT-ESTIV Webinar Series
Webinar 1, February 2024

Disclaimer: Inotiv staff provide technical support for NICEATM, but do not represent NIEHS, NTP, or the official positions of any federal agency.

National Institutes of Health • U.S. Department of Health and Human Services
Background

• Understanding how human populations interact with and are exposed to chemical sources is essential for contextualizing chemical hazard and understanding chemical risk
  – Many chemicals lack measured estimates of human exposure

• High-throughput exposure simulations and structure-based chemical use models can help inform exposure scenarios for data-poor chemicals
  – The results of high-throughput methods can be difficult to navigate, especially for those unfamiliar with computational methods

Adapted from Wambaugh et al. 2019, Current Opinion in Toxicology
In 2023, we integrated exposure predictions from the Environmental Protection Agency’s (EPA’s) Systematic Empirical Evaluation of Models (SEEM3) and functional use categories from EPA’s Chemical and Product Database (CPDat) into ICE.
SEEM3 was developed by the ExpoCast group at EPA’s Center for Computational Toxicology and Exposure.

Figure adapted from Ring et al. 2019 Env. Sci. Tech.
• We pulled SEEM3 predictions for over 600,000 chemicals from EPA’s github page (https://github.com/HumanExposure/SEEM3RPackage) in November 2022
  – 5\textsuperscript{th}, 50\textsuperscript{th}, and 95\textsuperscript{th} percentile of exposure in mg/kg/day
  – Limited dataset to \textasciitilde480,000 chemicals that were within the model’s domain of applicability
Exposure in ICE: In Vitro to In Vivo Extrapolation (IVIVE)

Adapted from Wambaugh et al. 2019, Current Opinion in Toxicology
SEEM3 Exposure Pathways

- Exposure pathway is often not known
- SEEM3 predicts pathway of exposure through chemical structure-based and property-based machine learning models
  - Consumer
  - Dietary
  - Far-field pesticides
  - Far-field industrial

Figure from Ring et al. 2019 Env. Sci. Tech.
• For the ~480,000 chemicals we also pulled the pathway predictions from SEEM3

• We created our near-field and far-field annotations based on the predicted pathway

Most chemicals have no predicted pathway!
Direct Access of Exposure Predictions Data Set

- ICE Data Sets page contains downloadable data set with SEEM3 exposure estimates, SEEM3 predicted pathways, and ICE-curated near-field and far-field annotations.

Download Data from ICE.
https://ice.ntp.niehs.nih.gov/DATASETDESCRIPTION
Chemical Use Categories in ICE

• Chemical use categories explain the various ways in which a chemical can be used
  – In what sectors is a chemical used? What role does the chemical play? What products a chemical can be found in? etc.

• Can help inform pathway for exposure models

• In ICE, our chemical use categories are derived from CPDat (v3 2021, Williams et al. 2017)

[Image of the EPA website]

https://www.epa.gov/chemical-research/chemical-and-products-database-cpdat
Curated Product Use Categories

- Derived from EPA’s Product Use Categories (Isaacs et al. 2019, JESEE)

- Curated product use categories describe the consumer products a chemical may be in
  - Useful in determining chemical composition, exposure frequency, and route of exposure

- Users can visualize and extract these categories from the ICE Chemical Characterization tool
  - Over 300 categories for nearly 5000 chemicals

Bubble plots show hierarchical structure and relative abundance of categories

Summary tables provide breakdown by category and chemical
Functional use is the role a chemical plays within a product. We pull two types of functional use from CPDat:

1. Reported functional use harmonized to Organisation of Economic Co-operation and Development (OECD) categories
2. Predicted functional use is predicted from Quantitative Structure Use Relationship (QSUR) models (Phillips et al. 2018)
OECD Functional Use

• Within CPDat, reported functional use for ~7,500 chemicals are harmonized to uses established by the OECD

• We harmonized reported use to OECD use for ~2,000 additional chemicals

• Functional use dataset in ICE has 77 OECD uses for 9,395 chemicals
Within CPDat, functional use predictions are already limited to chemicals within the QSUR model’s applicability domain. We curate the data further by limiting the probability of prediction to $\geq 80\%$ to ensure high confidence in results.

Functional use dataset in ICE has 37 predicted functional uses for 192,438 chemicals.
Visualizations of Functional Use Data

- Accessible through the ICE Chemical Characterization tool

The number of substances with the given type of functional use.

Color of the cell corresponds with chemical list in chemical query.
Direct Access to Functional Use Data Set

- Accessible through the ICE Datasets Page

**Chemical Use**

*Chemical and Products Database (CPDat)*

Curated product use categories, OECD functional use categories, and predicted functional use categories in ICE are derived from EPA’s CPDat database. CPDat is a document-based database that contains exposure-relevant data on chemicals in consumer products (Dionisi et al. 2018). It is actively updated and maintained by EPA’s Center for Computational Toxicology and Exposure. Data in ICE were curated from CPDat v3 (January 2021).

Curated product use categories and functional use mappings for chemicals in ICE can be explored via the Functional Use Explorer and Curated Product Use Explorer visualizations in the ICE Chemical Characterization tool.

<table>
<thead>
<tr>
<th>Chemical Name</th>
<th>CASRN</th>
<th>DTXSID</th>
<th>Endpoint</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,3-Cyclohexanediethanamine</td>
<td>2579-20-6</td>
<td>DTXSID4041238</td>
<td>OECD Functional Use</td>
<td>Solvent</td>
</tr>
<tr>
<td>1,3-Cyclohexanediethanone</td>
<td>2579-20-6</td>
<td>DTXSID4041238</td>
<td>OECD Functional Use</td>
<td>Processing aid not otherwise specified</td>
</tr>
<tr>
<td>3-Isothiazolone</td>
<td>1003-07-2</td>
<td>DTXSID9074935</td>
<td>OECD Functional Use</td>
<td>Adhesivo/cohesion promoter</td>
</tr>
<tr>
<td>3-Isothiazolone</td>
<td>1003-07-2</td>
<td>DTXSID9074935</td>
<td>OECD Functional Use</td>
<td>Antioxidant</td>
</tr>
<tr>
<td>4-Formylphenylboronic acid</td>
<td>87199-17-5</td>
<td>DTXSID6074887</td>
<td>OECD Functional Use</td>
<td>Ph regulating agent</td>
</tr>
<tr>
<td>4-Formylphenylboronic acid</td>
<td>87199-17-5</td>
<td>DTXSID6074887</td>
<td>OECD Functional Use</td>
<td>Biocide</td>
</tr>
<tr>
<td>D-Glucopyranoside, hexyl</td>
<td>54549-24-5</td>
<td>DTXSID1041843</td>
<td>OECD Functional Use</td>
<td>Surfactant</td>
</tr>
<tr>
<td>1,3-Butanediol, polymer with</td>
<td>68400-67-9</td>
<td>DTXSID90100952</td>
<td>OECD Functional Use</td>
<td>Adhesivo/cohesion promoter</td>
</tr>
</tbody>
</table>

*ICE Data Sets*

- Data Sets
- Acute Lethality
- Cancer
- Cardiotoxicity
- DART
- Endocrine
- Irritation-Corrosion
- Sensitization
- CHTS
- Chemical Parameters
- Exposure Predictions
- Chemical Use
Future Additions: ClassyFire Chemical Taxonomies

- ClassyFire (Djoumbou et al. 2016, Wishart Research Group) is an automated, structure-based, hierarchical chemical taxonomy with up to 11 levels of classification and 4,285 classifications across all levels of the hierarchy.
Future Additions: ClassyFire Chemical Taxonomies

- ClassyFire (Djoumbou et al. 2016, Wishart Research Group) is an automated, structure-based, hierarchical chemical taxonomy with up to 11 levels of classification and 4,285 classifications across all levels of the hierarchy.

- One potential use case for these classifications is to link certain chemical use categories within ICE to ClassyFire taxonomies.
  - Can help identify chemical classes and structures that are most abundant in their chemical query.

- Identifying chemical groups of interest can help focus follow-up investigations or aid in the selection of alternative chemicals.
We conducted a case study of 100 chemicals with the OECD functional use of biocides to demonstrate a potential ClassyFire output, which mapped to 88 chemicals within the ClassyFire Kingdom "Organic Compounds”

<table>
<thead>
<tr>
<th>Superclass</th>
<th>Class</th>
<th># Chemicals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benzenoids</td>
<td>Benzene and substituted derivatives</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Naphthalenes</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Phenol ethers</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Phenols</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Triphenyl compounds</td>
<td>1</td>
</tr>
<tr>
<td>Lipids and lipid-like molecules</td>
<td>Fatty Acyls</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Glycerolipids</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Prenol lipids</td>
<td>3</td>
</tr>
<tr>
<td>Organic acids and derivatives</td>
<td>Carboxylic acids and derivatives</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Organic carbonic acids and derivatives</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Organic sulfuric acids and derivatives</td>
<td>1</td>
</tr>
<tr>
<td>Organic nitrogen compounds</td>
<td>Organonitrogen compounds</td>
<td>10</td>
</tr>
<tr>
<td>Organic oxygen compounds</td>
<td>Organooxygen compounds</td>
<td>9</td>
</tr>
<tr>
<td>Organoheterocyclic compounds</td>
<td>Azoles</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Azolidines</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Benzothiazoles</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Diazinanes</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Isoindoles and derivatives</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Metallo heterocyclic compounds</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Oxazinanes</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Pyridines and derivatives</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Quinolines and derivatives</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Triazinanes</td>
<td>1</td>
</tr>
<tr>
<td>Organosulfur compounds</td>
<td>Thioethers</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Thioureas</td>
<td>1</td>
</tr>
<tr>
<td>Phenylpropanoids and polyketides</td>
<td>Cinnamic acids and derivatives</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Depsides and depsidones</td>
<td>1</td>
</tr>
</tbody>
</table>
We conducted a case study of 100 chemicals with the OECD functional use of biocides to demonstrate a potential ClassyFire output, which mapped to 88 chemicals within the ClassyFire Kingdom "Organic Compounds"
Summary and Future Directions

• In our continuing efforts to provide high-confidence, high-quality toxicologically relevant data, we curated exposure predictions from EPA’s SEEM3 model and functional use data from EPA’s CPDat
  – This data was made publicly available with the release of ICE v4.0 (March 2023) and ICE v4.0.1 (August 2023)

• With the inclusion of these data, users will be able to better explore how human populations may interact with chemicals and their potential levels of exposure

• The addition of these new data types into ICE facilitates the potential addition of new data sources, exposure models, and types of use
  – Occupational exposure models, models with demographic information, other types of use like sector of use, etc.