2019 Jan 15;53(2):719-732. doi: 10.1021/acs.est.8b04056. Epub 2018 Dec 24.

National Institute of Environmental Health Sciences
Division of Translational Toxicology

Curating Chemical Use Categories and Exposure Predictions to Inform Chemical Assessment

Victoria Hull[1], Tripp LaPratt[1], Agnes L. Karmaus[1*], Kimberly T. To[1], Aswani Unnikrishnan[1], David G. Allen[1], Nicole C. Kleinstreuer[2]

[1]Inotiv, Research Triangle Park, NC; [2]NIH/NIEHS/DTT/NICEATM, Research Triangle Park, NC

Introduction

- Understanding how human populations interact with and are exposed to chemical sources is essential to prioritizing chemicals for risk assessment, as risk is the intersection of chemical hazard and chemical exposure (Wambaugh et al. 2019).
  - However, many chemicals lack measured estimates of human exposure and little is known about how they are used.
- High-throughput exposure simulations and structure-based chemical-use models can help inform exposure scenarios for data-poor chemicals.
  - The results of these high-throughput models can be difficult to navigate, especially for those unfamiliar with computational methods.
- The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) has curated chemical exposure predictions from the U.S. Environmental Protection Agency's (EPA) Systematic Empirical Evaluation of Models (SEEM3, Ring et al. 2019) and functional use categories from EPA's Chemicals and Products Database (CPDat, Williams 2017) to help place chemical hazard in the larger context of chemical risk.
  - Functional use data is the role a chemical plays within a product. Both measured and predicted uses were curated for inclusion in ICE.
- We integrated these exposure predictions and new use categories into NICEATM's Integrated Chemical Environment (ICE; https://ice.ntp.niehs.nih.gov/) which is an open-access resource containing toxicologically relevant data and computational tools.

Source of Exposure Predictions

- SEEM3 was developed by the ExpoCast group at EPA's Center for Computational Toxicology and Exposure (Figure 1A).
  - This model predicts both population-level exposure in mg/kg/day for over 600,000 chemicals and the potential pathway of exposure (Figure 1B).

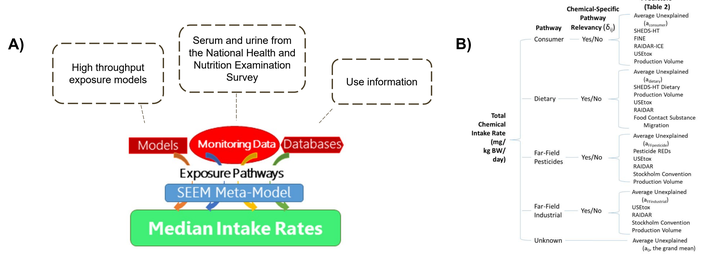

Figure 1: Figure is adapted from **Ring et al. 2019**. A) An overview of the SEEM3 model inputs and outputs. B) The pathways predicted by the SEEM3 model and the high-throughput exposure models that are used for each pathway.

Pathways of Exposure for Chemicals in ICE

- SEEM3 predictions were pulled for over 600,000 chemicals from EPA's GitHub page in November 2022 (https://github.com/HumanExposure/SEEM3RPackage). Outputs of the model include the 5th, 50th, and 95th percentile of exposure in mg/kg/day.
  - The data set was further refined to ~480,000 chemicals that were within the model's domain of applicability (Figure 2).

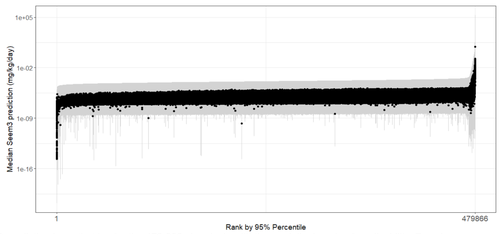

Figure 2: The Median SEEM3 prediction in mg/kg/day for the 479,866 chemicals with the model's domain of applicability. Error bars represent the 5th and 95th percentile of exposure. Chemicals are ranked by 95% percentile.

- For the ~480,000 chemicals within the domain of applicability we also pulled the pathway predictions from SEEM3 and created near-field and far-field annotations based on the predicted pathway (Figure 3).
  - Over 80% of chemicals do not have a known pathway, as chemical use information is often limited.

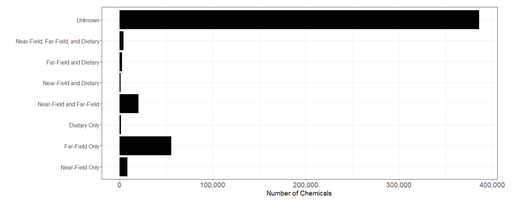

Figure 3: The number of chemicals with a given pathway of exposure.

Visualize New Exposure Data in ICE

- Exposure predictions can be overlayed onto boxplots of equivalent administered dose (EAD) in the ICE In Vitro to In Vivo Extrapolation (IVIVE) tool (Figure 4).
- This feature allows users to see the intersection of chemical hazard (as characterized by EAD) and predicted exposure, providing a more comprehensive picture for potential chemical risk.

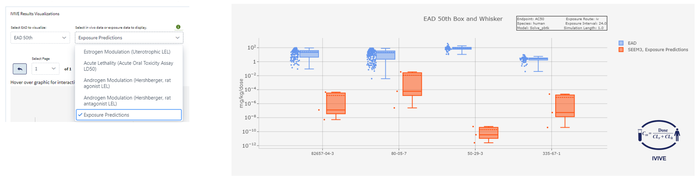

Figure 4: An example output of the ICE IVIVE tool with boxplots showing EAD and SEEM3 exposure predictions.

Chemical Use Data in ICE

- To help users explore potential chemical use and exposure scenarios, ICE contains chemical use categories that are derived from EPA's CPDat v3.
  - CPDat is a document-based database that contains exposure-relevant data on chemicals in consumer products (Dionisio et al. 2018).
- Curated product use categories describe the products that chemicals are reported to be in.
  - Within ICE there are over 300 different product use categories for 4896 unique chemicals.
  - Knowing product use can help determine factors like chemical composition, exposure frequency, and route of exposure.
  - Users can visualize and extract these categories from the ICE Chemical Characterization tool (Figure 5).

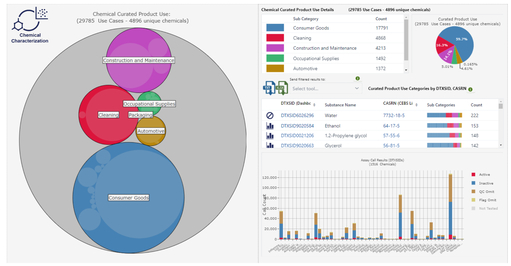

Figure 5: Bubble plot and summary table depicting curated product use categories in the ICE Chemical Characterization tool.

Curation of Functional Use Data

- Functional use can help inform what products a chemical is used in if product use category is not known, how much of a chemical is used within a product, and potential chemical alternatives.
- Two types of functional use were pulled from CPDat, including reported functional use harmonized to Organization of Economic Co-operation and Development (OECD) categories and predicted functional as predicted by quantitative structure–use relationship models (QSUR, Phillips et al. 2017).
  - As part of the curation process, reported use was additionally harmonized to OECD use for ~2000 chemicals.
  - For predicted functional use, the prediction probability was limited to ≥ 80% to ensure high-confidence results.
- The functional use dataset in ICE has 77 OECD uses for 9,395 chemical (Figure 6A) and 37 QSUR-predicted uses for 192,438 chemicals (Figure 6B).

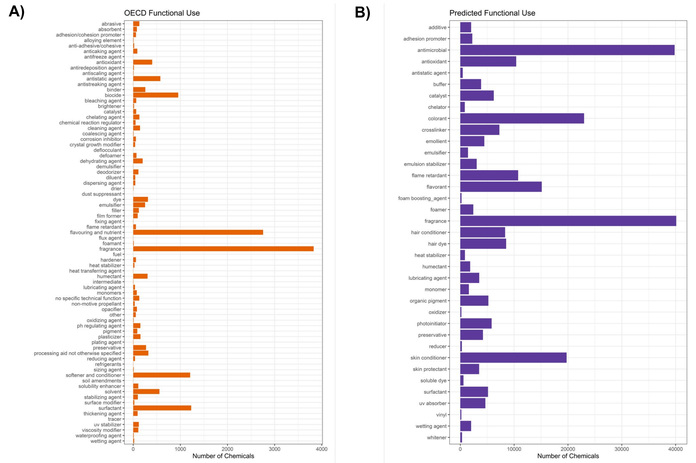

Figure 6: A) The number of chemicals that fall within each OECD functional category in the ICE functional use dataset. B) The number of chemicals that fall within each QSUR-predicted functional use category in the ICE functional use dataset.

Visualization of New Functional Use Data in ICE

- ICE users can visualize and extract OECD functional use and predicted functional use via heatmaps in the ICE Chemical Characterization tool (Figure 7).

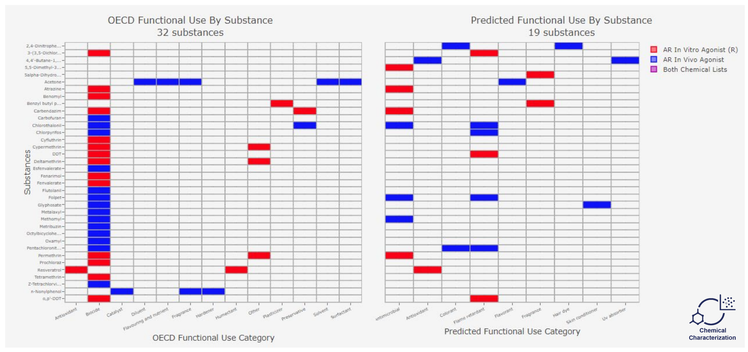

Figure 7: An example heatmap to visualize functional use categories within the ICE chemical characterization tool. Heatmap was generated using a query for two chemical lists, including the AR In Vitro Agonist and AR In Vivo Agonist Chemical Quick Lists in ICE. Color of the cell corresponds with chemical list. Plot titles indicate the amount of substances queried with a given type of use, as some chemicals may have predicted use only, OECD use only, or both OECD use and predicted use.

Accessing Data

- Exposure predictions and functional use data can be downloaded directly through the ICE data sets page (https://ice.ntp.niehs.nih.gov/DATASETDESCRIPTION).

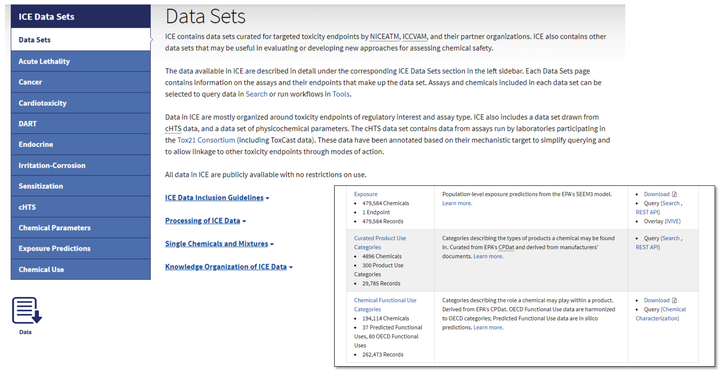

Figure 8: A screenshot of the ICE data sets page where users can download data and access descriptions of data sources and curation processes.

Future Additions: ClassyFire Chemical Taxonomies

- ClassyFire (Djoumbou et al. 2016) is an automated, structure-based, hierarchical chemical taxonomy created by the Wishart Research group. It has up to 11 levels of classification with 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, allowing ICE users to 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 (Figure 9).

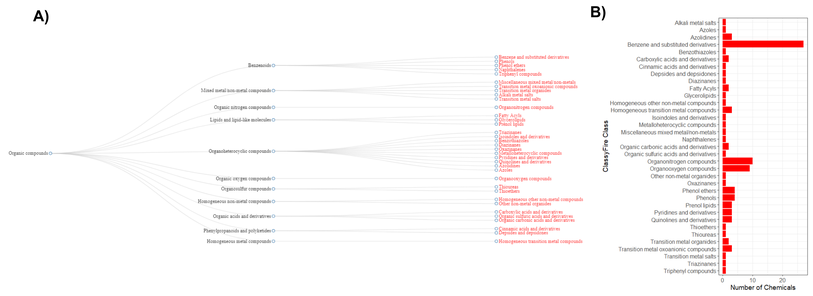

Figure 9: ClassyFire results for 100 chemicals identified as biocides in ICE. A) A taxonomic tree for the Kingdom, SuperClass, and Class levels of the ClassyFire hierarchy. B) The number of chemicals in each class for the "Class" level of the hierarchy.

Summary and Discussion

- 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 includes exposure predictions for over 475,000 chemicals, OECD functional use categories for nearly 10,000 chemicals, and predicted functional use data for approximately 192,000 chemicals.
- These new data are publicly available in ICE with releases v4.0 (March 2023) and v4.0.1 (August 2023). Data can be directly downloaded from the ICE data sets page or visualized in the IVIVE and Chemical Characterization tools.
  - With the inclusion of these data, users will be able to better explore how human populations may interact with chemicals and potential levels of exposure.
- In future releases of ICE, ClassyFire chemical taxonomies will be incorporated to allow users to see how chemical structures can be connected with different exposure scenarios and chemical use cases.
- The addition of exposure and use data into ICE facilitates the potential addition of new data sources, exposure models, and use types.

References

Dionisio et al. 2018 Sci. Data 10;5:180125. https://doi.org/10.1038/sdata.2018.125.
Djoumbou et al. 2016. J. Cheminfo. 8:16. https://doi.org/10.1186/s13321-016-0174-y.
Phillips et al. 2017. Green Chem. 19(4): 1063–1074. https://doi.org/10.1021/acs.est.8b04056.
Ring et al. 2019. Environ Sci Technol. 15;53(2):719-732. https://doi.org/10.1021/acs.est.8b04056.
Wambaugh et al. 2019. Current Opinion in Toxicology 15:76-92. https://doi.org/10.1016/j.cotox.2019.07.001.
Williams. 2017. The Chemical and Products Database (CPDat) MySQL Data File. The United States Environmental Protection Agency's Center for Computational Toxicology and Exposure. Dataset. https://doi.org/10.23645/epacomptox.5352997.

Acknowledgments

The Intramural Research Program of the National Institute of Environmental Health Sciences (NIEHS) supported this poster. Technical support was provided by Inotiv under NIEHS contract HHSN273201500010C. Since the poster was written as part of the official duties of the authors, it can be freely copied. We thank Catherine Sprankle and Barbara Stevenson, Inotiv, for editorial input.

*A.L. Karmaus is currently affiliated with Syngenta, Greensboro, NC

Visit ICE
https://ice.ntp.niehs.nih.gov/