Curating chemical use categories and exposure predictions to inform chemical assessment

V. Hull[1], T.R. LaPratt[1], A.L. Karmaus[1], K.T. To[1], A. Unnikrishnan[1], D.G. Allen[1], N.C. Kleinstreuer[2]

[1]Inotiv, RTP, NC, United States;[2]NIH/NIEHS/DTT/NICEATM, RTP, NC, United States

Background and Purpose

Understanding how human populations interact with and are exposed to chemicals is essential for prioritizing chemicals for risk assessment, as risk is the intersection of chemical hazard and exposure. High-throughput in silico exposure simulations and chemical use models can inform exposure scenarios for data-poor chemicals. However, the large data volumes associated with these tasks can be challenging for those unfamiliar with computational methods. To provide easily interpretable and accessible exposure and use data, we integrated exposure predictions from the U.S. Environmental Protection Agency's (EPA) Systematic Empirical Evaluation of Models (SEEM3) and use categories from EPA's Chemical and Product Database (CPDat) into the Integrated Chemical Environment (ICE; https://ice.ntp.niehs.nih.gov/). ICE is an open-access resource developed by the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) to provide highly curated toxicologically relevant data and approachable computational tools. Through ICE, users will be able to extract, visualize, and explore these new data sources.

Methods

To compile a robust exposure dataset, we pulled SEEM3 predictions for over 600,000 chemicals from the EPA GitHub page (https://github.com/HumanExposure/SEEM3RPackage). The extracted outputs include the 5th, median, and 95th percentiles of exposure in mg/kg/day, and exposure pathway predictions for dietary, consumer, far-field industrial, far-field pesticide, and unknown pathways. To curate the data we refined the data set to chemicals that were within the model's applicability domain and we used predicted pathways to generate near-field and far-field annotations. To obtain use data, we extracted two types of functional use categories from CPDat v3: reported functional use harmonized to Organization for Economic Co-operation and Development (OECD) categories, and functional use as predicted by EPA quantitative structure–use relationship (QSUR) models. Functional use categories explain the function that chemicals serve in a product, such as a solvent, abrasive, or humectant. As part of the curation process, we additionally harmonized reported use to OECD use for ~2000 chemicals. For predicted functional use, the prediction probability was limited to ≥ 80%, a threshold we defined as high-confidence results.

Results

Exposure predictions for 480,000 chemicals within the SEEM3 model applicability domain were curated and included in ICE. These exposure predictions, their predicted pathway, and the near- and far-field annotations can be downloaded directly through the ICE data sets page. The curated SEEM3 exposure predictions were also integrated into the ICE In Vitro to In Vivo Extrapolation (IVIVE) tool where users can overlay the 5th, median, and 95th percentiles of predicted exposure onto boxplots of equivalent administered doses. This allows users to translate chemical hazard and bioactivity into predicted margins of exposure, providing a more comprehensive visualization of potential chemical hazard vs exposure. The curated functional use dataset

contains 77 OECD uses for 9,395 chemicals and 37 QSUR-predicted uses for 192,438 chemicals. To access functional use data, ICE users can either download the data through the ICE data sets page or visualize OECD functional use and predicted functional use with heatmaps in the ICE Chemical Characterization tool.

Conclusions

Presented alongside other toxicologically relevant data, the addition of exposure predictions and functional use categories into ICE provides users with added context for evaluating chemicals. The addition of exposure and use data into ICE facilitates the future addition of new data sources, exposure models, and use types. This work will be expanded with ongoing efforts to curate and integrate chemical structure classifications using the ClassyFire chemical taxonomy from the Wishart Research Group, thereby allowing users to examine associations between chemical structure with different use and exposure scenarios. These taxonomies can assist with identifying chemical groups of interest and can help focus follow-up investigations or aid in the selection of alternative chemicals. This project was funded in whole or in part with federal funds from the NIEHS, NIH under Contract No. HHSN273201500010C.

Keywords: Computational Toxicology, Chemical Characterization, Exposure Assessment