Probabilistic Points of Departure: A Tiered Approach for Life Cycle Impact Assessment

Workshop: Advancing Quantitative Analysis in Human Health Assessments through Probabilistic Methods

8-Oct-2024

Weihsueh A. Chiu, PhD
Texas A&M University

Major Collaborators: Nicolo Aurisano, Kerstin Johanna Felicitas von Borries, Peter Fantke, Jacob Kvasnicka, Olivier Jolliet

Acknowledgments

Final WHO/IPCS Author Group

- Bernard Bottex, EFSA representative
- David Bussard, U.S. EPA
- Weihsueh Chiu, formerly U.S. EPA
- George Fotakis, ECHA representative
- Andy Hart, FERA, UK
- Dale Hattis, Clark University, USA
- Matthias Herzler, BfR, Germany
- Kathy Hughes, IPCS
- Wout Slob, RIVM, Netherlands
- Theo Vermeire, RIVM, Netherlands
- Carolyn Vickers, IPCS

Other colleagues and trainees

- ***Nicolo Aurisano, PhD (Maersk)**
- ***Kerstin von Borries, PhD (DTU)**
- ***Peter Fantke, PhD (USEtox)**
- Suji Jang, PhD, (trainee, now at ExxonMobil)
- **Olivier Jolliet, PhD (DTU)**
- **Richard Judson, PhD (EPA, retired)**
- ***Jacob Kvasnicka, PhD (TAMU, now at EPA)**
- Hsing-Chieh Lin, PhD, TAMU
- En-Hsuan Lu, PhD, (trainee, starting at MN PCA)
- Greg Paoli, RSI
- Ivan Rusyn, MD PhD, TAMU
- Kan Shao, Indiana U
- Lauren Zeise, California EPA

NIH/NIEHS P42 ES027704
NIH/NIEHS P30 ES029067
NIH/NIEHS T32 ES026568

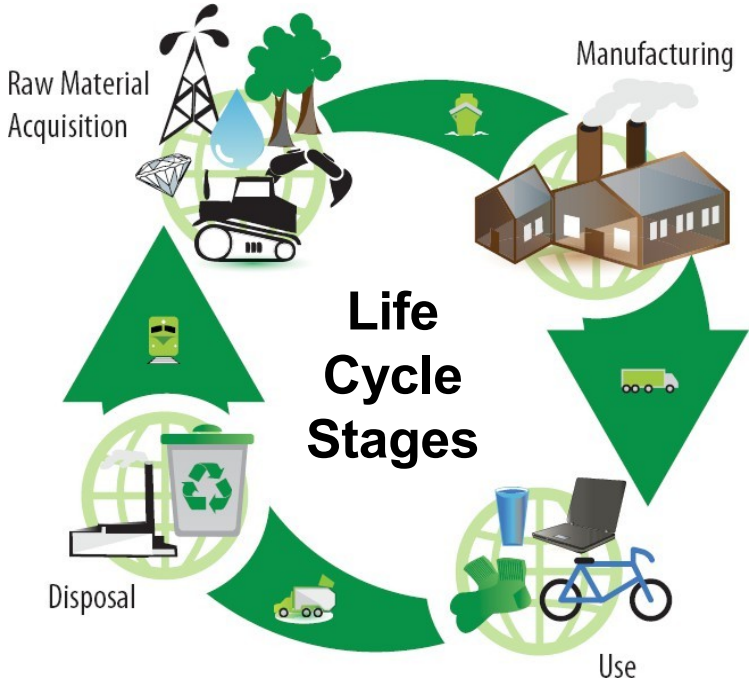
U.S. EPA STAR RD84004601
U.S. EPA STAR RD83580201

Additional Funding from 'Global Best Practices on Emerging Chemical Policy Issues of Concern under UN Environment's Strategic Approach to International Chemicals Management (SAICM)' (GEF project ID 9771, grant no. S1-32GFL-000632)

***Special thanks for providing slides from which this presentation is adapted**

What is the Life Cycle?

"the major activities in the course of the product's life-span from its manufacture, use, and maintenance, to its final disposal, including the raw material acquisition required to manufacture the product." -EPA 2006 (Life Cycle Assessment: Principles and Practice)



Often iterative in nature

Goal and Scope Definition

Life Cycle Inventory

Life Cycle Impact Assessment

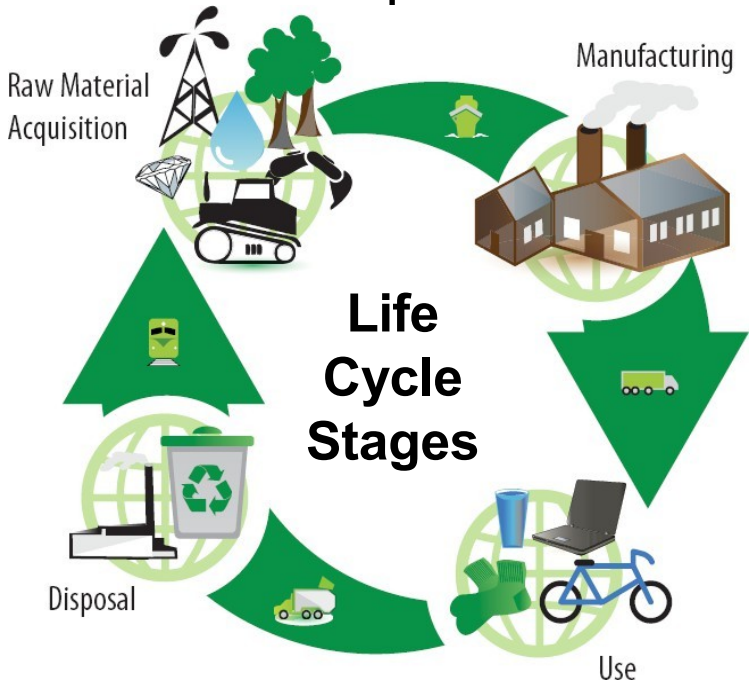
Interpretation

Impacts on Health and Environment

Inputs

- Raw materials
- Energy

Etc.



Outputs

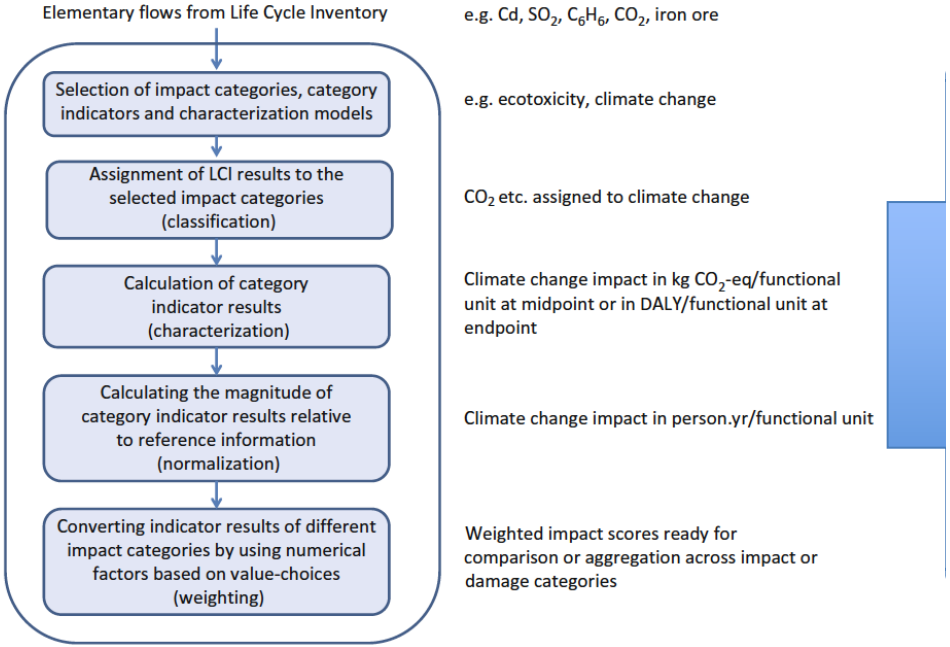
- Atmospheric emissions
- Waterborne wastes
- Solid wastes
- Co-products

Etc.

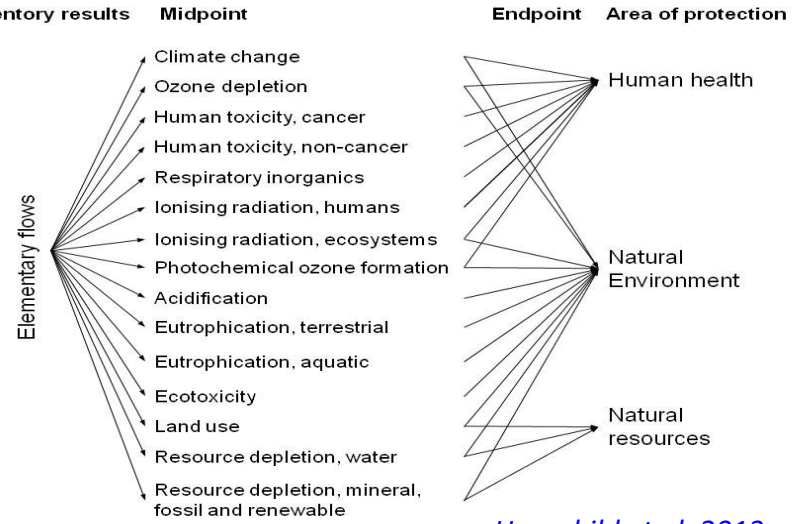
Impacts on Health and Environment

What does this have to do with Toxicology?

Five Steps of LCIA



Framework for Impact Categories



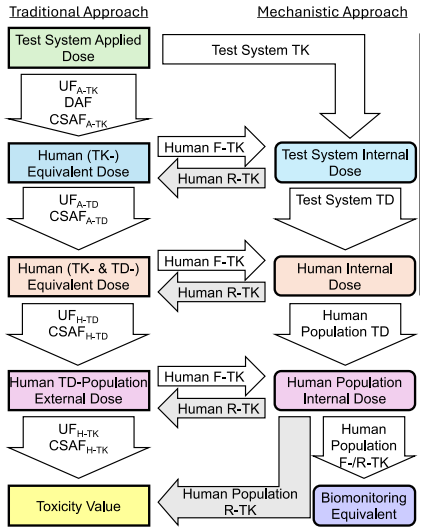
[Hauschild et al. 2013]

ISO 14040 standard, adapted by Hauschild and Huijbregts (2015) *Life Cycle Impact Assessment*, Ch 1

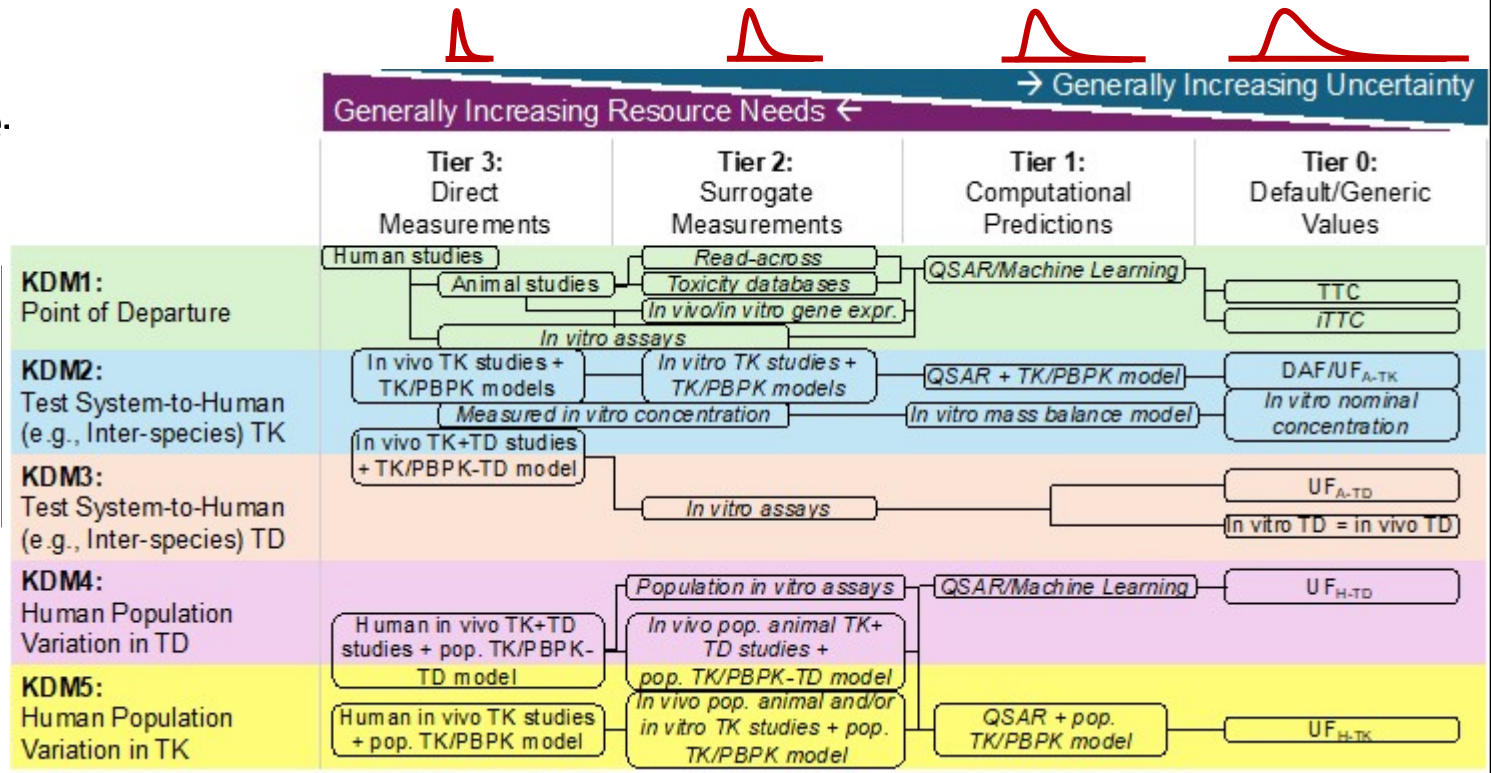
Context for this work

- 30,000−100,000 unique chemical substances are commonly used worldwide in various products, processes, or services
- LCIA is a comparative assessment approach that includes characterizing toxicological impacts on human health from all possible chemical exposures associated with the life cycles of those products, processes, or services
- Points of Departure (PODs) are are essential part of characterizing toxicity-related human health impacts in LCA
- Regulatory/authoritative PODs cover a very limited set of chemicals
- Treating “no number” chemicals as non-toxic underestimates impacts, which can bias decision-making

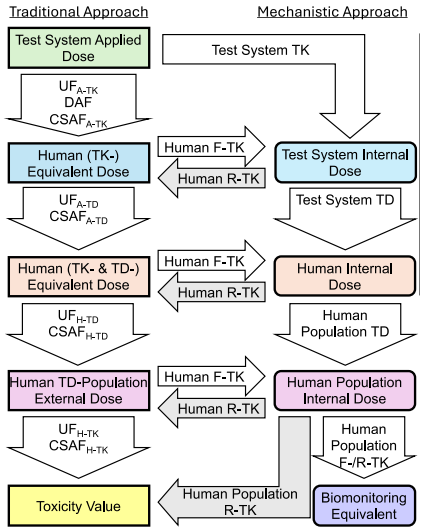
Unified Probabilistic Framework for Dose-Assessment



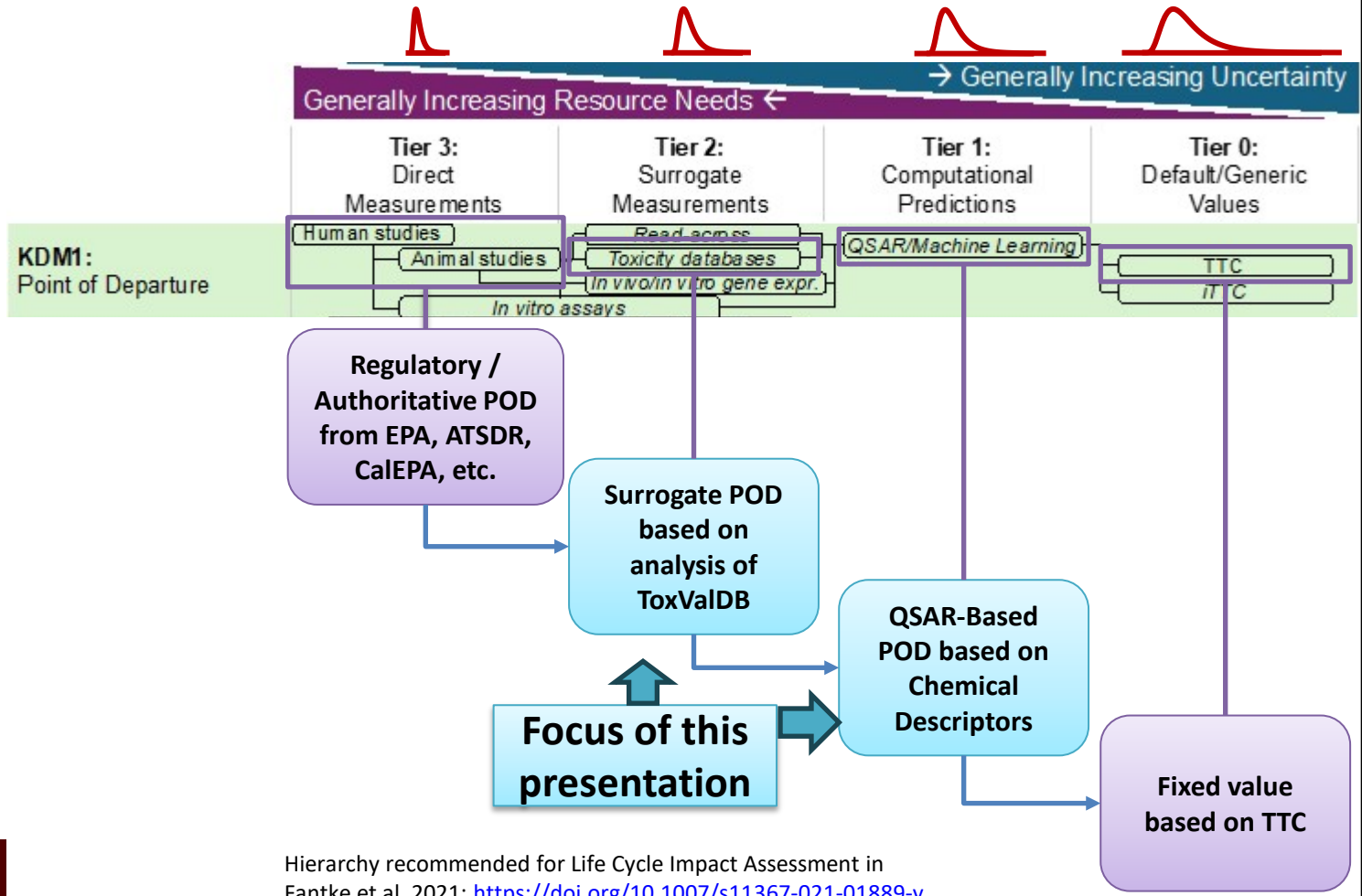
Lu et al. (in press, JTEH-B)



Focus on KDM 1: Point of Departure Determination



Lu et al. (in press, JTEH-B)



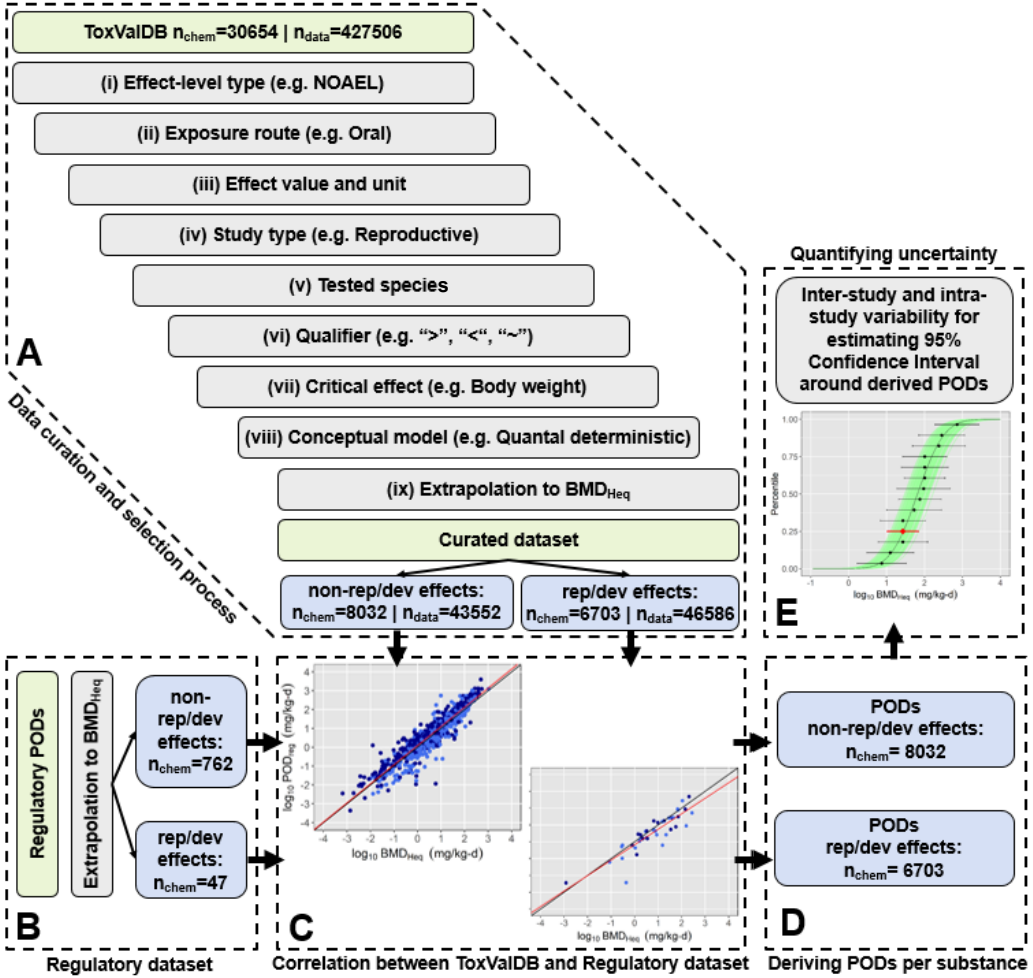
Hierarchy recommended for Life Cycle Impact Assessment in Fantke et al. 2021: https://doi.org/10.1007/s11367-021-01889-y

Surrogate POD based on analysis of ToxValDB

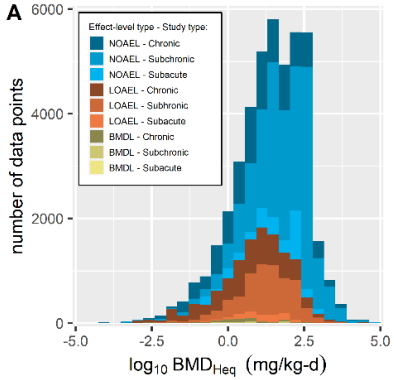
Develop a workflow for deriving PODs <u>with quantified uncertainty</u> for chemical substances <u>with</u> animal toxicology data but <u>without</u> regulatory/authoritative assessments.

A: Data curation and selection
B-C: Calibration to overlapping regulatory PODs from authoritative sources
D: Application to dataset from **A**
E: Uncertainty analysis

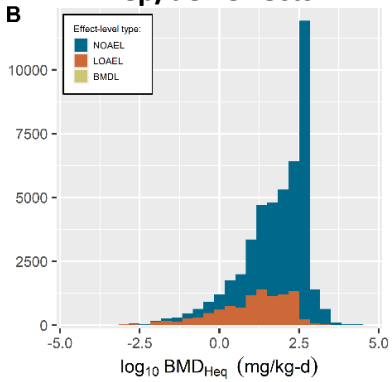
Oral PODs: Aurisano et al. 2023
Inhalation PODs: Aurisano et al. 2024

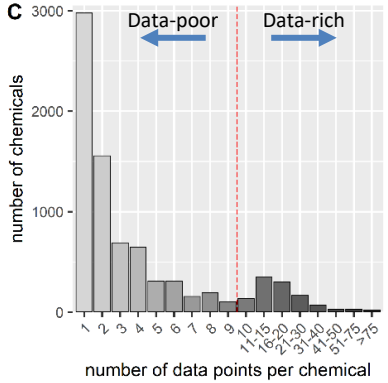


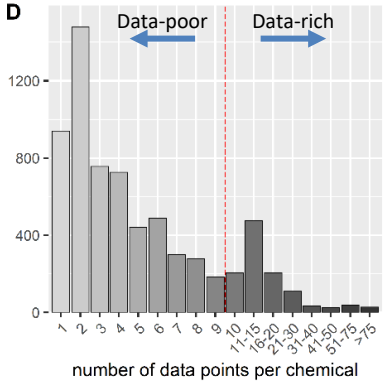
General non-cancer effects

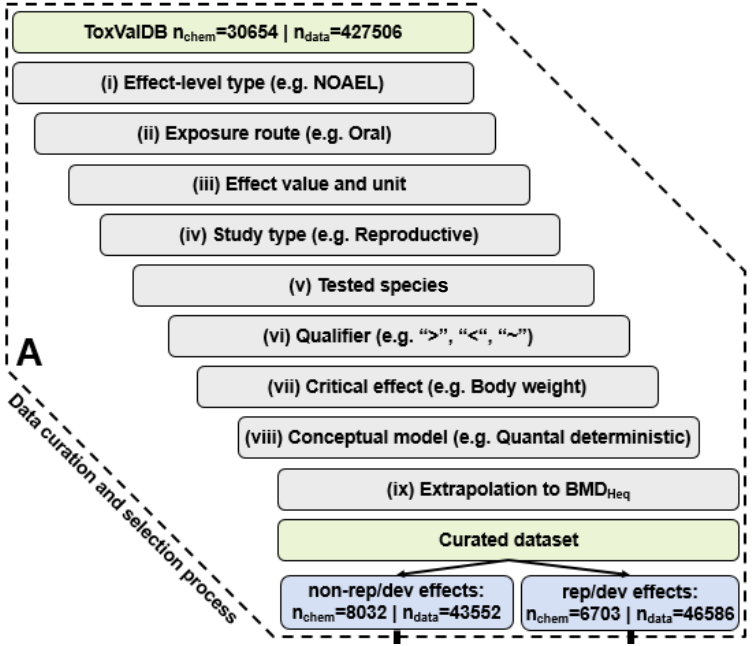


Rep/dev effects

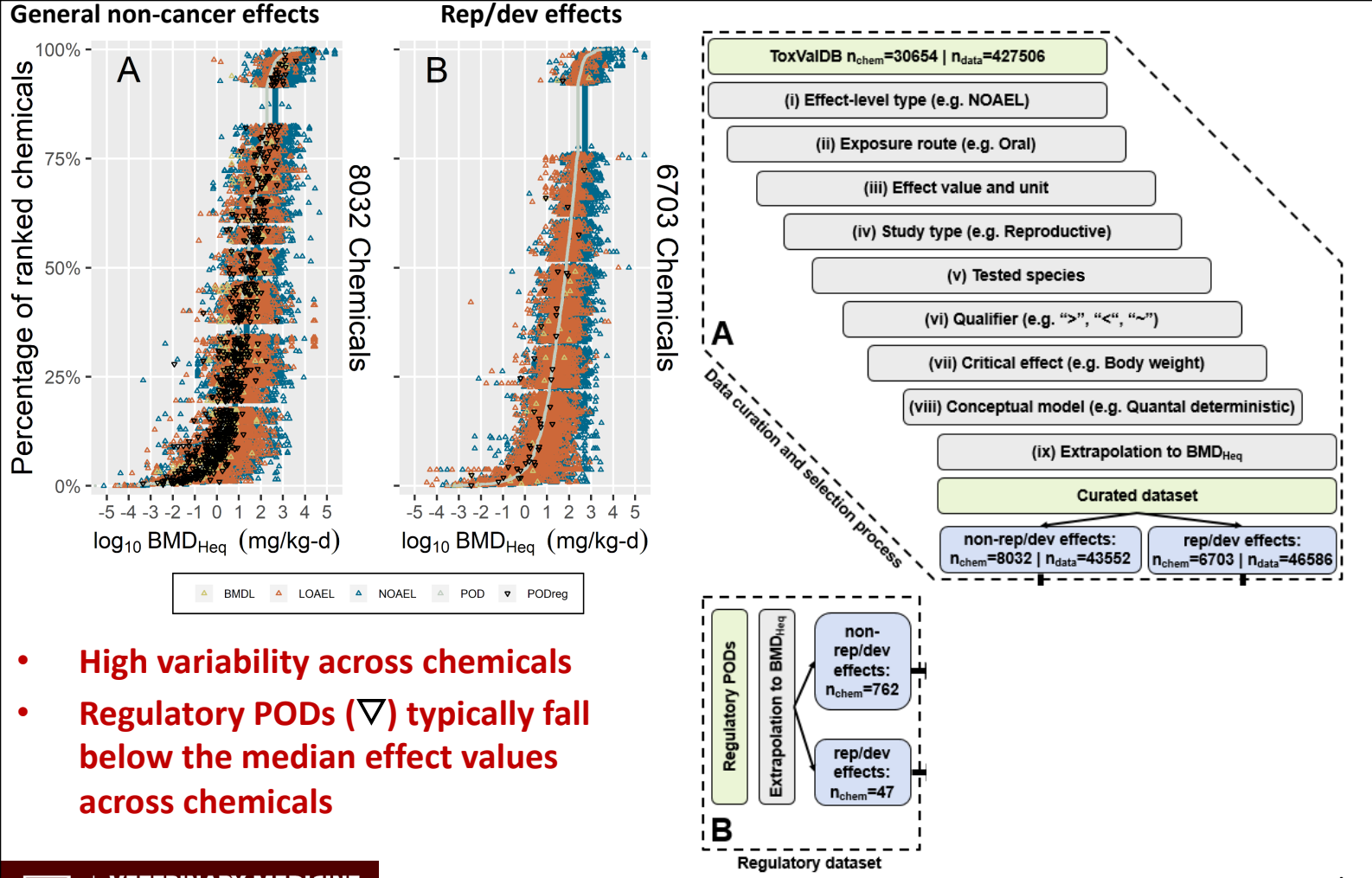








Aurisano et al. 2023, 2024



- **High variability across chemicals**
- **Regulatory PODs (▽) typically fall below the median effect values across chemicals**

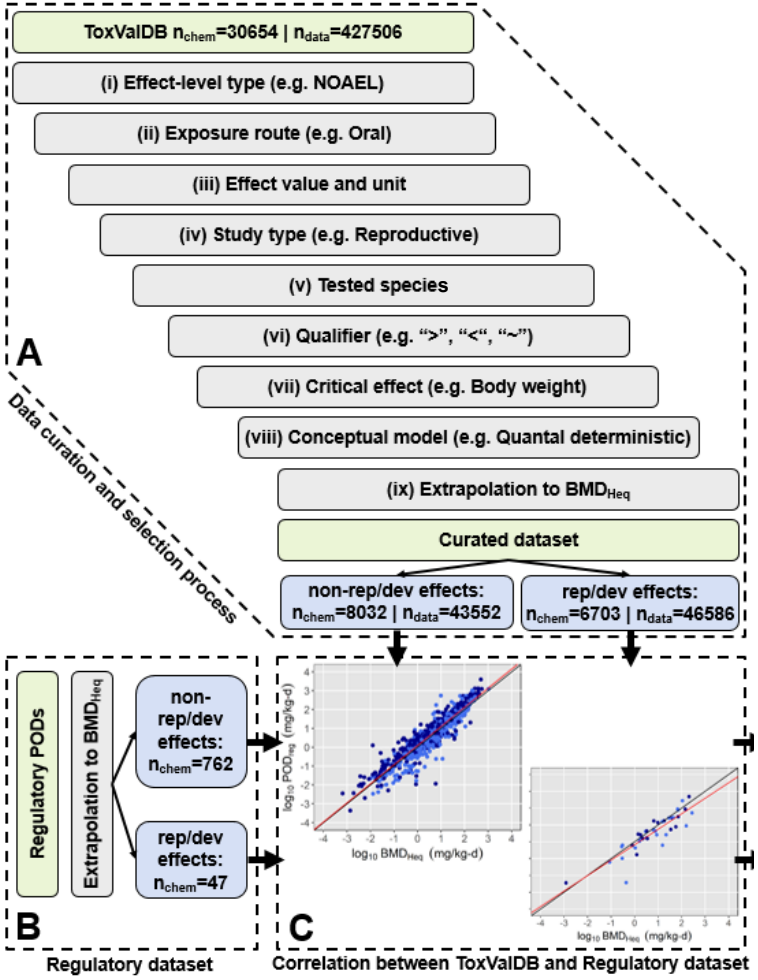
Aurisano et al. [2023](), [2024]()

Simple hypothesis: Regulatory PODs can be "modeled" as a "conservative" %ile of the (curated) ToxValDB data for each chemical.

- **Data-rich:** %ile from fitted on a log-normal distribution
- **Data-poor:** %ile from a fixed log-normal distribution ("avg" data-rich chemical)

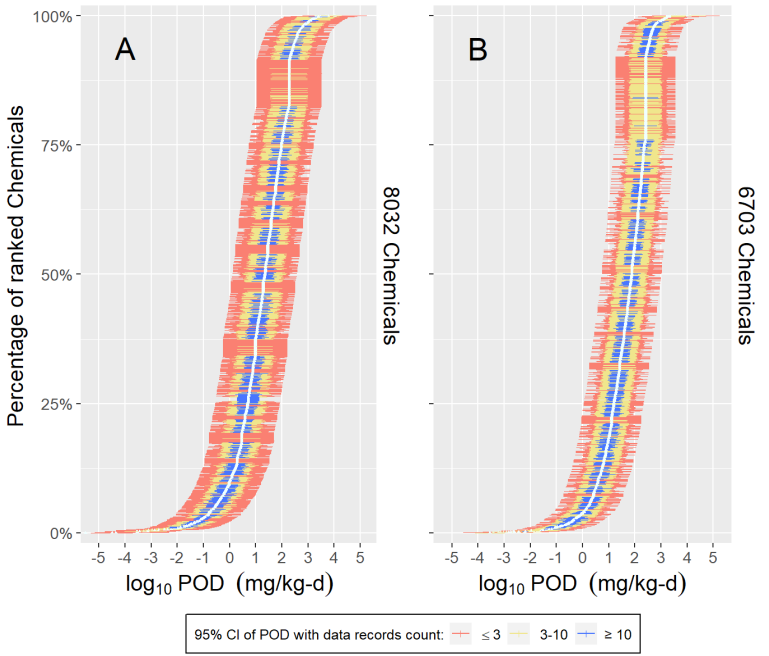
Results using 25th %-ile

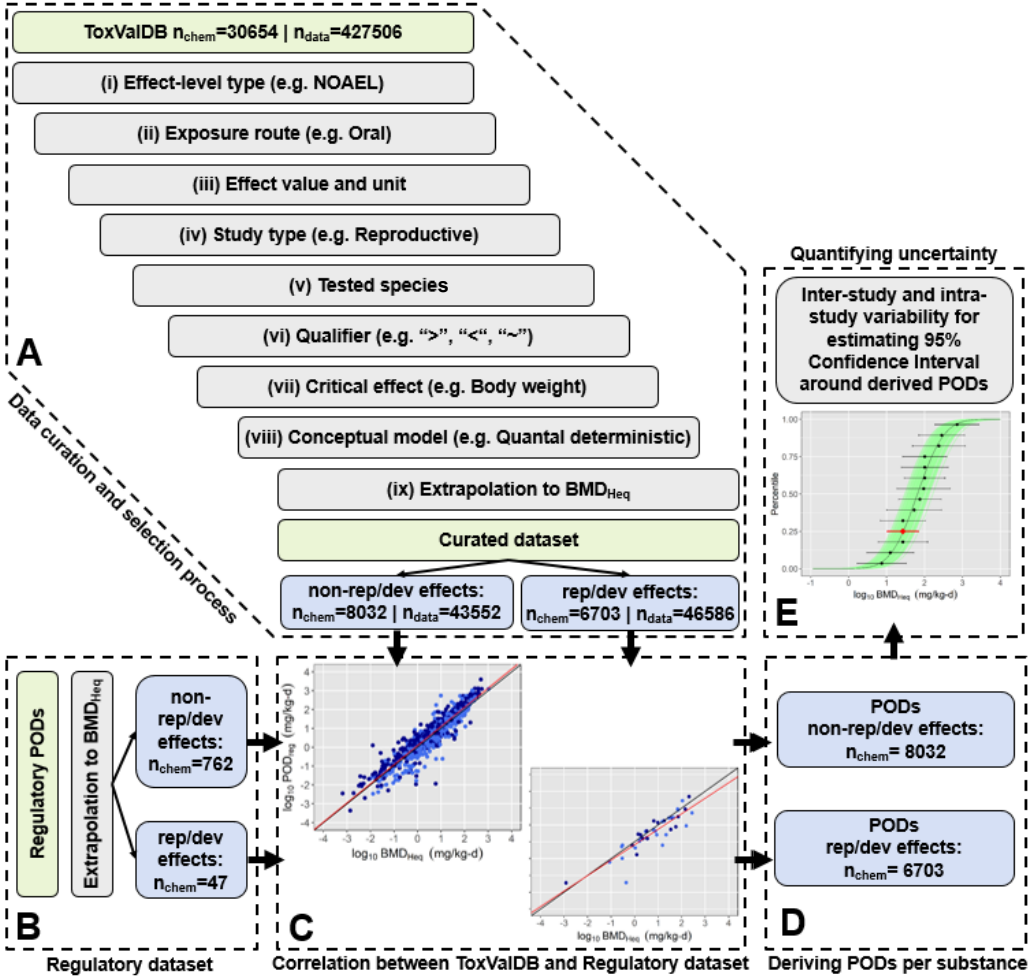
- Oral General non-cancer (n=744):
  R^2=0.85 RSE = 0.46 (log10 units)
- Oral Rep/dev (n=41):
  R^2=0.78 RSE = 0.53 (log10 units)
- Inhalation General & Rep/dev (n=174):
  R^2=0.76 RSE = 0.82 (log10 units)



Aurisano et al. 2023, 2024

Tiered uncertainty characterization based on number of (curated) ToxValDB data points, inter-, and intra-study variability





Aurisano et al. 2023, 2024

Surrogate POD based on analysis of ToxValDB

Approach: Expand coverage of chemicals with (non-cancer) toxicity values by

- Created a consistent and curated data set of *in vivo* chronic dose-response toxicity data from EPA ToxValDB
- Developed a statistical approach for calibrating toxicity data against regulatory values
- Quantified uncertainty from inter- and intra-study variability

Results: Surrogate PODs can be derived using the 25th %ile from ToxValDB

- Oral PODs expanded by ***n* > 10,000**
- Inhalation PODs expanded by ***n* > 2,000**

Limitations

- Tens of thousands of chemicals have no or inadequate data in ToxValDB
- *In vivo* testing data on these chemicals unlikely to expand substantially in the near future

Machine Learning to the Rescue?

Aurisano et al. 2023, 2024

QSAR-Based POD based on Chemical Descriptors

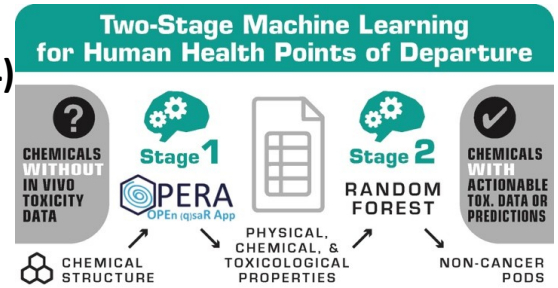
Conditional Toxicity Value (CTV) Predictor (2018)

- QSAR built on regulatory toxicity values
- Predicts oral and inhalation (experimental) NOAELs

Two-Stage Machine Learning Model (2024)

- QSAR for PODs building on surrogate oral PODs from Aurisano et al. (2023)
- Model for inhalation PODs in development

Both approaches perform better than ToxCast/in vitro NAMs for predicting regulatory PODs



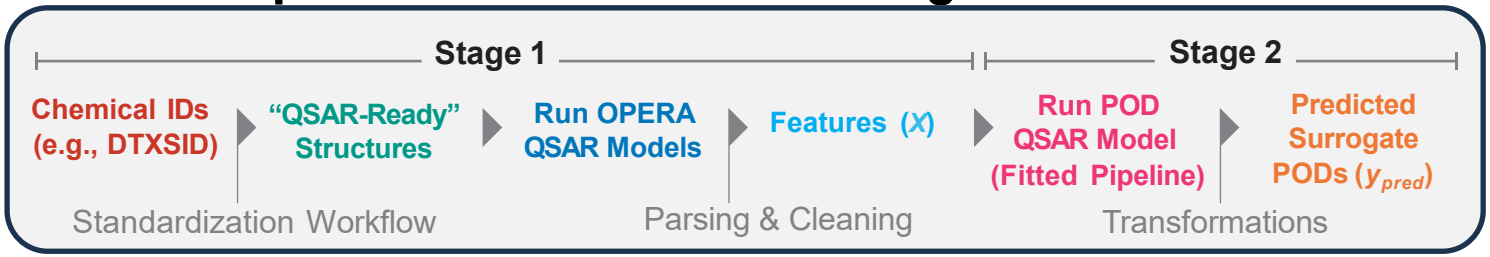
https://wchiu.shinyapps.io/Two-Stage-ML-Results-Browser/

Approach	RMSE	MedAE	R^2
CTV	N.R.	0.70	0.45
Two-Stage ML (general non-cancer)	0.69	0.40	0.48
Two-Stage ML (repro/dev)	0.58	0.31	0.49
ToxCast+httk (general non-cancer)	1.87	1.22	<0
ToxCast+httk (repro/dev)	1.52	0.84	<0

RMSE: Root-mean-squared-error (log10 units)
MedAE: Median absolute error (log10 units)
R^2: Coefficient of determination (<0 means worse that naïve constant model)

Wignall et al. 2018; Kvasnicka et al. 2024

Conceptual Framework: Two-Stage QSAR Model

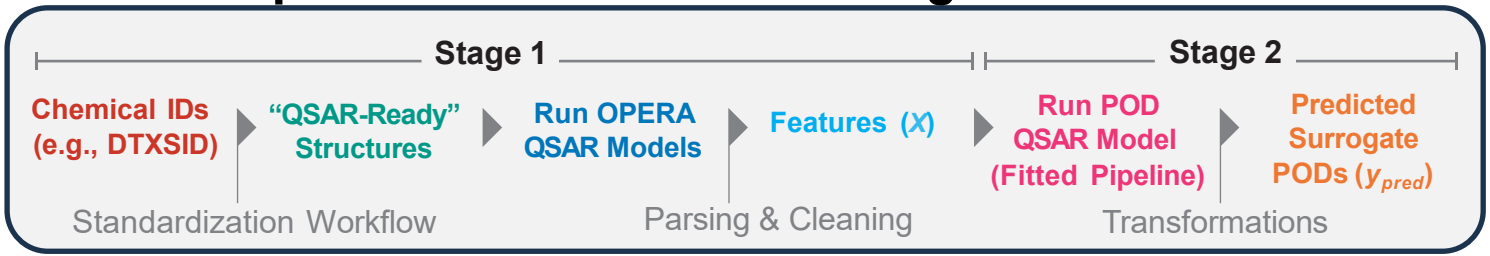


Why a two-stage model?

- Most chemical descriptors can be hard to interpret by a toxicologist or risk assessor (as opposed to a chemo-informaticist)
- Existing OPERA models provide open-source predictions for ***interpretable*** physical-chemical-toxicological parameters
- Analogous to a "supervised" neural network with a single intermediate layer composed of interpretable features.

Kvasnicka et al. 2024

Conceptual Framework: Two-Stage QSAR Model



Training Data Collection, & Preprocessing

g: general noncancer
rd: reproductive/developmental

Data Collection
Surrogate PODs from Aurisano et al. 2023 (y_{obs})

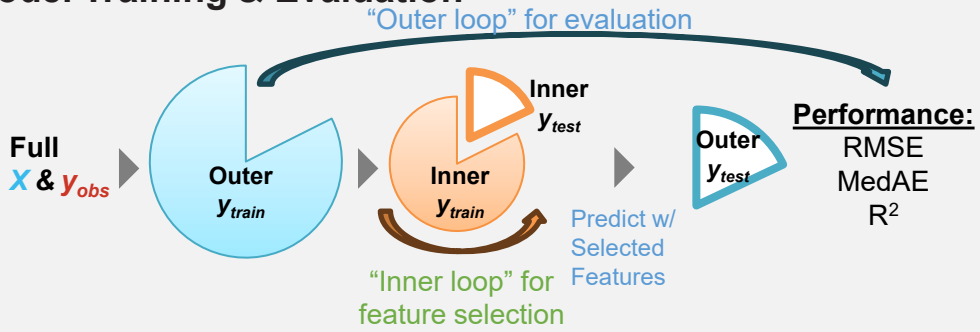
n_g = 5,209
n_{rd} = 4,938

Data Filtering
> 3 *in vivo* studies in ToxValDB & "QSAR-Ready"

n_g = 1,791
n_{rd} = 2,228

Feature Preparation
Run OPERA to generate features (X)

Model Training & Evaluation



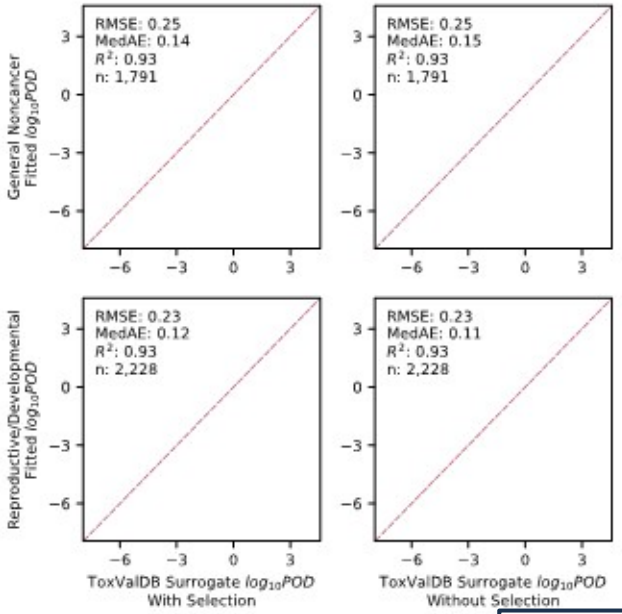
Performance:
RMSE
MedAE
R^2

Model Pipeline for Each Replicate
1. Feature Preprocessing
2. Random Forest Regression

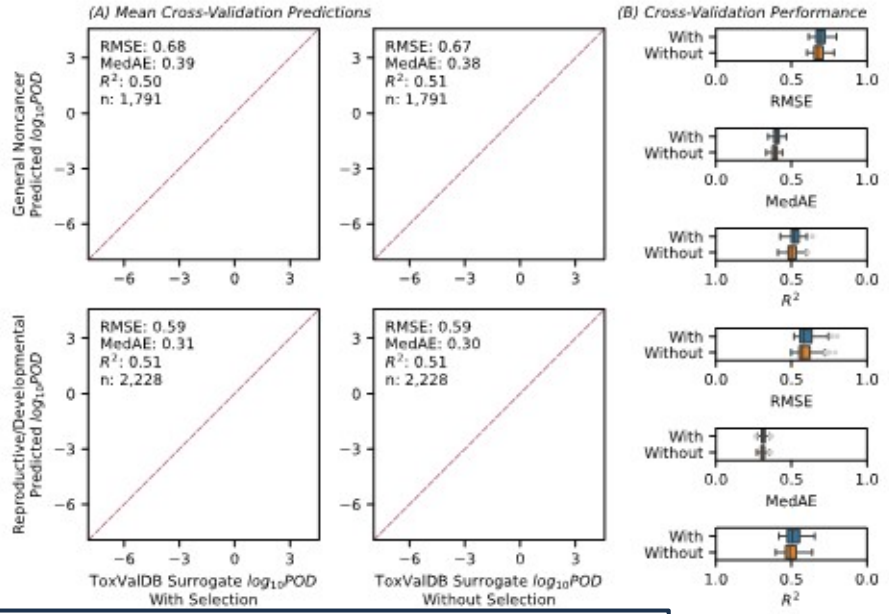
Kvasnicka et al. 2024

Model Evaluation

In-Sample Model Fitting



Out-of-Sample (Cross-Validation) Performance



Expected performance based on cross-validation results:

- Average Error (RMSE): **factor of 4~5**
- Typical Error (MedAE): **factor of 2~2.5**
- Explained Variance: ~**50%**

QSAR-Based POD based on Chemical Descriptors

https://wchiu.shinyapps.io/Two-Stage-ML-Results-Browser/

Applied to 800K+ chemicals from EPA CompTox Dashboard

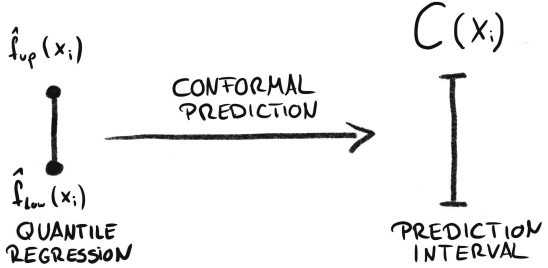
Limitations

- Same uncertainty estimate for every prediction
- Certain classes of chemicals excluded based on OPERA QSAR standardization workflow

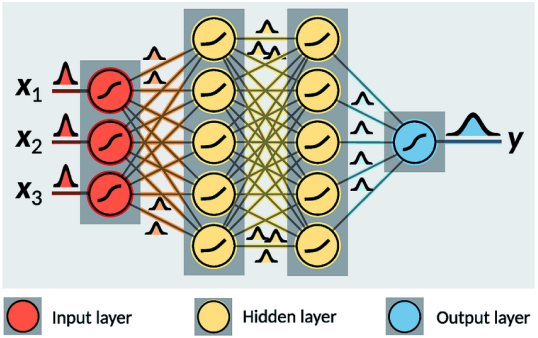
QSAR-Based POD based on Chemical Descriptors

Explore use of "Uncertainty-Aware" ML methods

- **Conformal prediction (CP)**



- **Bayesian neural network (BNN)**



- Addresses aleatoric *and* epistemic uncertainty
- Applied to >130K marketed chemicals

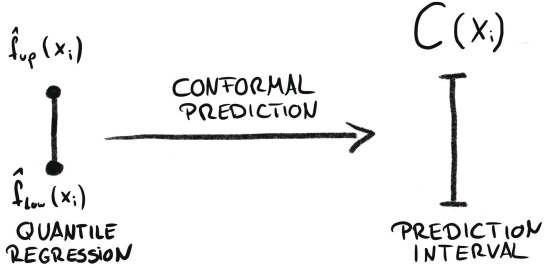
Preliminary Results

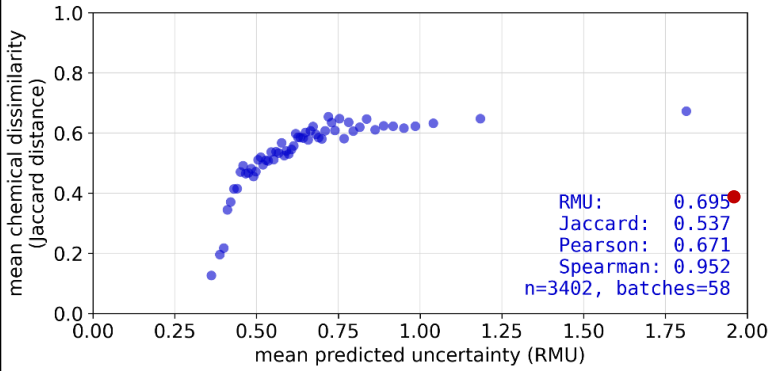
von Borries et al. in preparation

QSAR-Based POD based on Chemical Descriptors

Explore use of "Uncertainty-Aware" ML methods

- **Conformal prediction (CP)**





CP models performed better than BNN

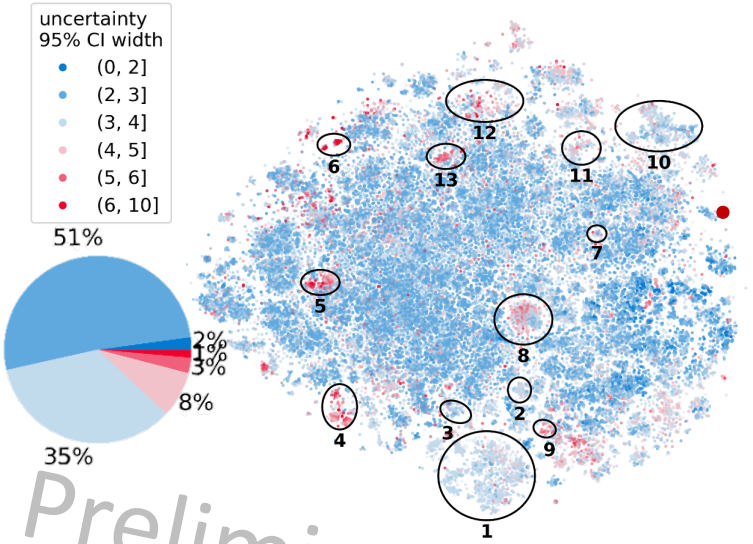
- Good coverage & well-calibrated confidence intervals
- Capture overall heteroscedasticity in prediction errors
- Higher uncertainty for new chemicals that are unlike training set chemicals

Preliminary Results

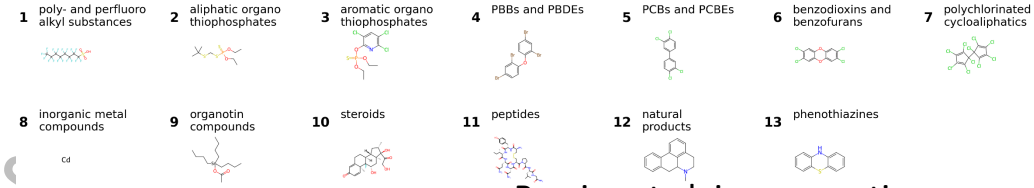
von Borries et al. in preparation

QSAR-Based POD based on Chemical Descriptors

Explore use of "Uncertainty-Aware" ML methods



- **Uncertainty hotspots**
  - polychlorinated and polybrominated compounds
  - metals and organometallics
  - alkaloids and phenothiazines
  - Peptides
- **Drivers of uncertainty**
  - low representation in the training data
  - low applicability of molecular descriptor developed for small organic molecules
  - (highly toxic) outliers



Preliminary Results

von Borries et al. in preparation

Summary: A Tiered Hierarchy of Probabilistic PODs

- LCIA requires PODs for tens-hundreds of thousands of chemicals for characterizing human health impacts of product, process, or service life cycles
- Regulatory/authoritative PODs cover a very limited set of chemicals
- Two classes approaches can fill these data gaps while also quantifying their uncertainty with varying degrees of sophistication
- **Key limitation:** Calibrated to existing regulatory PODs, which are largely based on experimental animal studies.
- **Counterfactual:** If we were to have new regulatory/authoritative assessments based on animal studies in the absence of any human data, we would make decisions based on them!