Two-Stage Machine Learning-Based Approach to Predict Points of Departure

Workshop: Advancing Environmental Health Research with Artificial Intelligence and Machine Learning

4-Nov-2024

Jacob Kvasnicka, PhD

Texas A&M University (now at U.S. EPA)

Major Collaborators: Weihsueh A. Chiu, Nicolò Aurisano, Kerstin von Borries, En-Hsuan Lu, Peter Fantke, Olivier Jolliet, Fred A. Wright

Disclaimer: This presentation does not necessarily reflect Agency policy.

Acknowledgments

Major Collaborators:

- ***Weihsueh A. Chiu, PhD (TAMU, Supervisor)**
- Nicolò Aurisano, PhD (Maersk)
- Kerstin von Borries, PhD (DTU)
- En-Hsuan Lu, PhD, (trainee, starting at MN PCA)
- Peter Fantke, PhD (USEtox)
- Fred A. Wright, PhD (NC State)
- Olivier Jolliet, PhD (DTU)
- Katherine A. Phillips, PhD (EPA)
- Kristin K. Isaacs, PhD (EPA)
- Peter Egeghy, PhD (EPA)

Other Support:

- Cedric Wannaz, PhD (MathWorks)
- Kamel Mansouri, PhD (NIEHS)
- Jian Tao, PhD (TAMU)
- ***Caroline Ring, PhD (EPA)**

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

Additional Funding from Swedish Foundation for Strategic Environmental Research (Grant No. DIA 2018/11), and the PARC project (Grant No. 101057014)

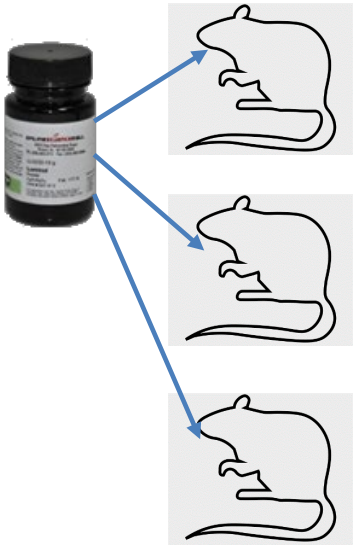
***Special thanks for providing slides from which this presentation is adapted**

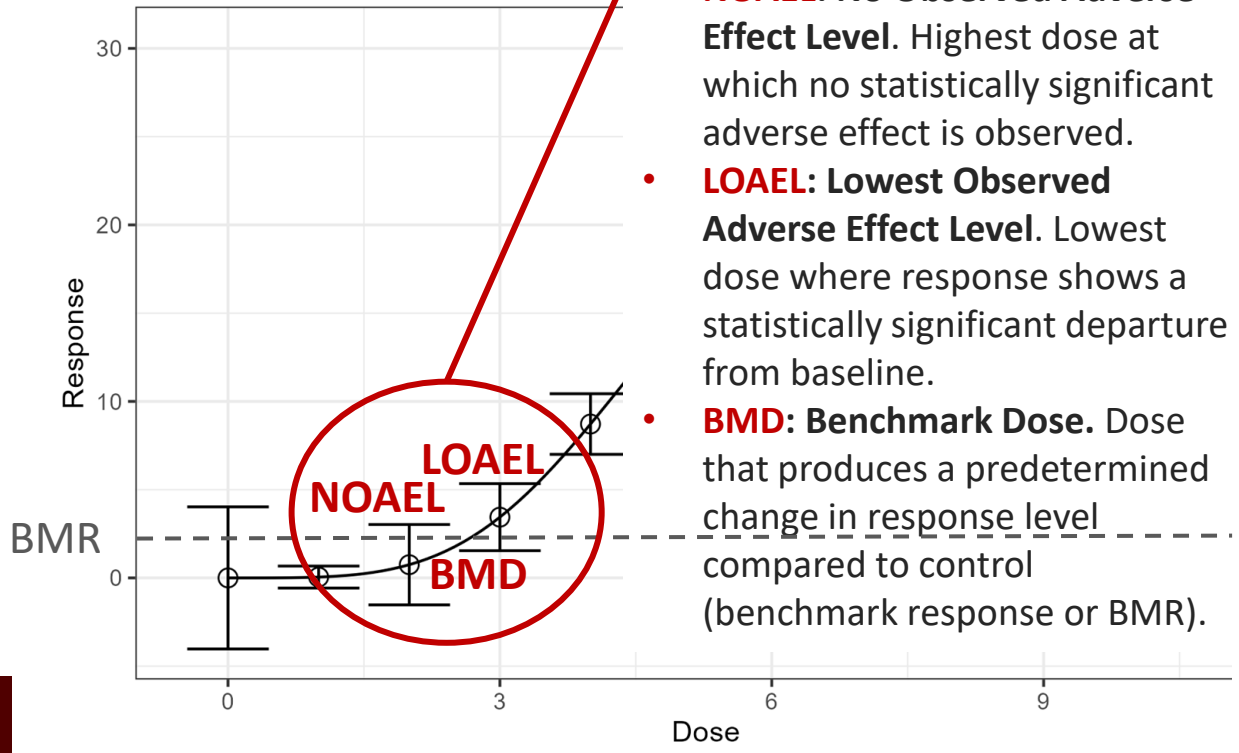
Context

- 30,000−100,000 unique chemicals used worldwide in various products, processes, or services
- **Points of Departure (PODs)** are essential for characterizing toxicity and assessing human health risks and impacts
- Regulatory/authoritative PODs cover a very limited set of chemicals
- Hypothesis: **Machine learning** can substantially expand the coverage of chemicals with **actionable PODs**

What is a POD? Toxicity Metric Derived from Experimental Dose-Response Study *In Vivo*

1. Dose groups of animals at different levels
2. Measure the response in each animal group
3. Determine the dose-response relationship

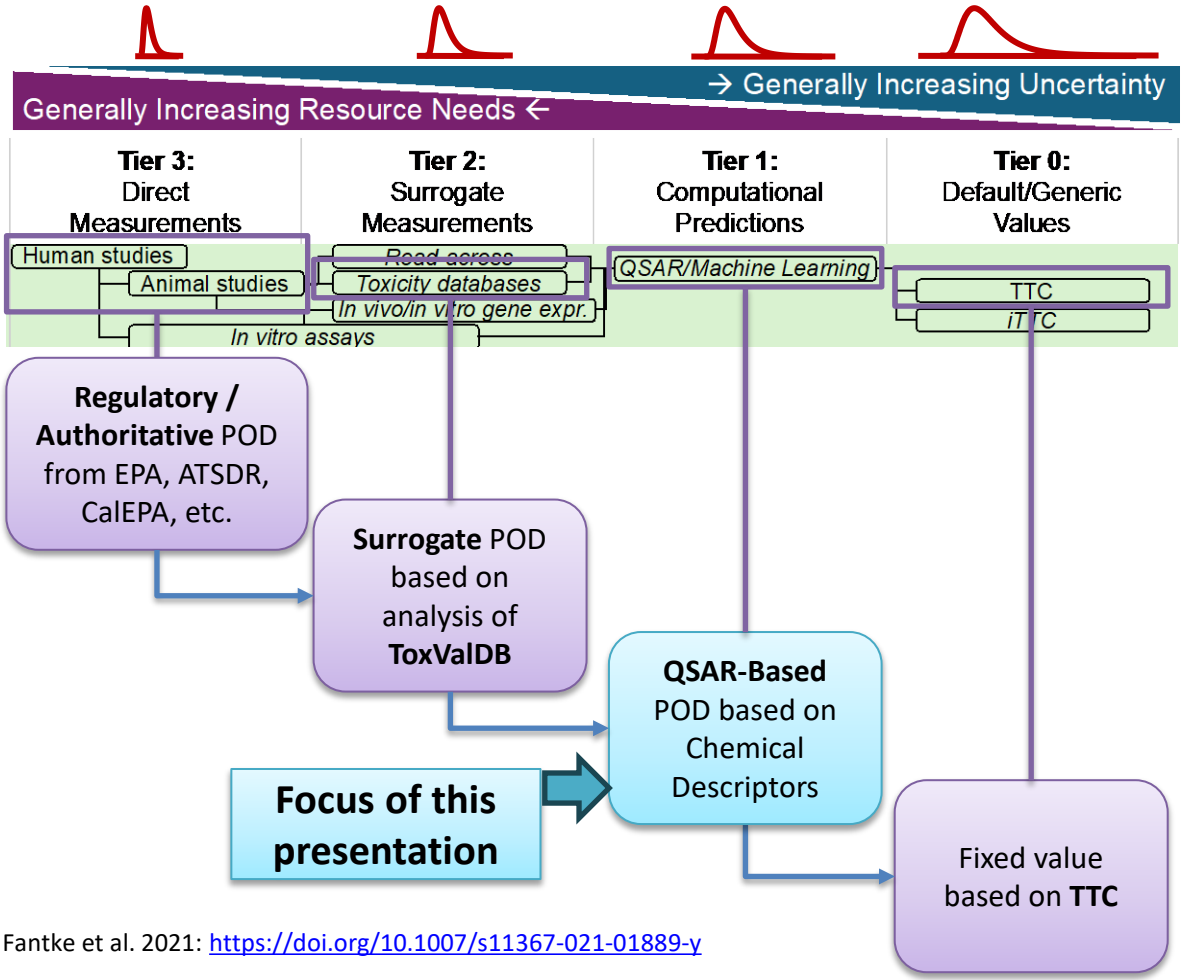




POD: Dose at which a significant departure from baseline response begins, indicating potential toxicity.

- **NOAEL: No Observed Adverse Effect Level**. Highest dose at which no statistically significant adverse effect is observed.
- **LOAEL: Lowest Observed Adverse Effect Level**. Lowest dose where response shows a statistically significant departure from baseline.
- **BMD: Benchmark Dose.** Dose that produces a predetermined change in response level compared to control (benchmark response or BMR).

Hierarchy of Approaches for Deriving a POD for Human Health Risk/Impact Assessment



Key Challenges

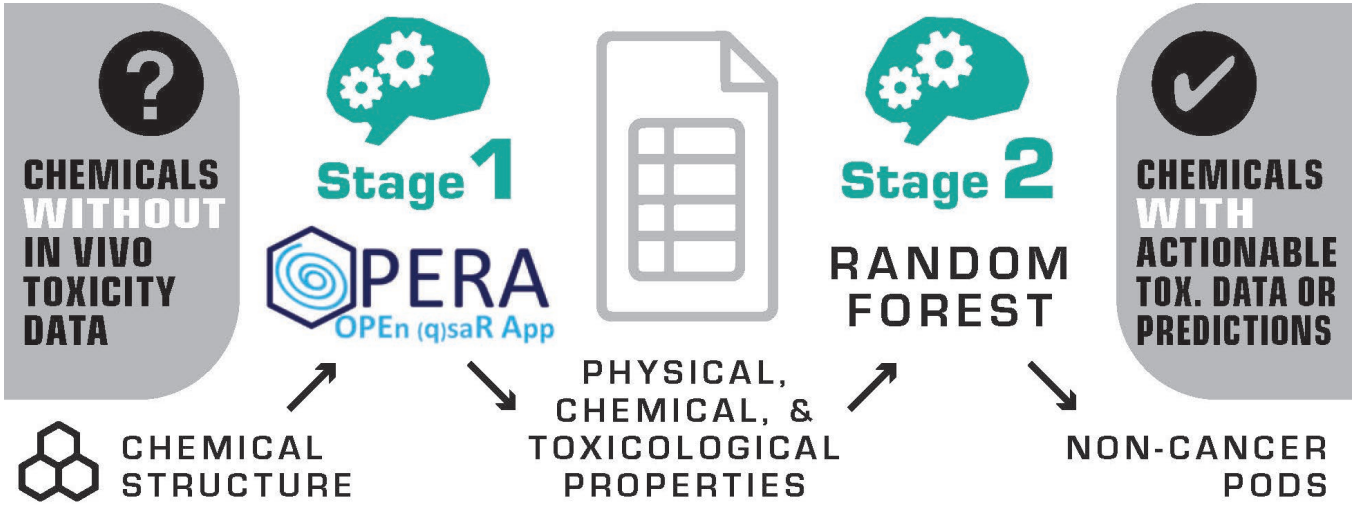
- Regulatory/authoritative PODs cover only several hundred chemicals
- Tens of thousands of chemicals have no or inadequate data in ToxValDB
- *In vivo* testing of these chemicals unlikely to expand substantially

Machine Learning to the Rescue?

Fantke et al. 2021: https://doi.org/10.1007/s11367-021-01889-y

Approach: Two-Stage QSAR Model Framework for Predicting Points of Departure

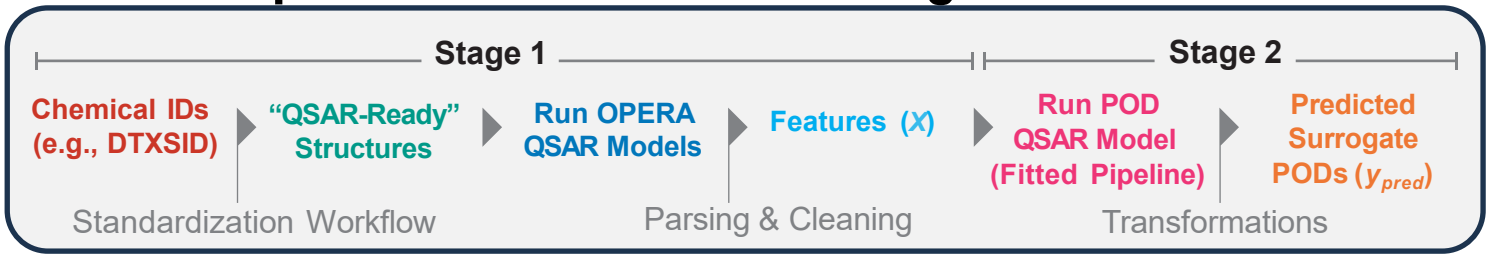
QSAR (Quantitative Structure-Activity Relationship): Uses Machine Learning to Predict Toxicity Based on Chemical Structure



OPERA: Mansouri et al. (2018-2024)

Kvasnicka et al. 2024

Conceptual Framework: Two-Stage QSAR Model

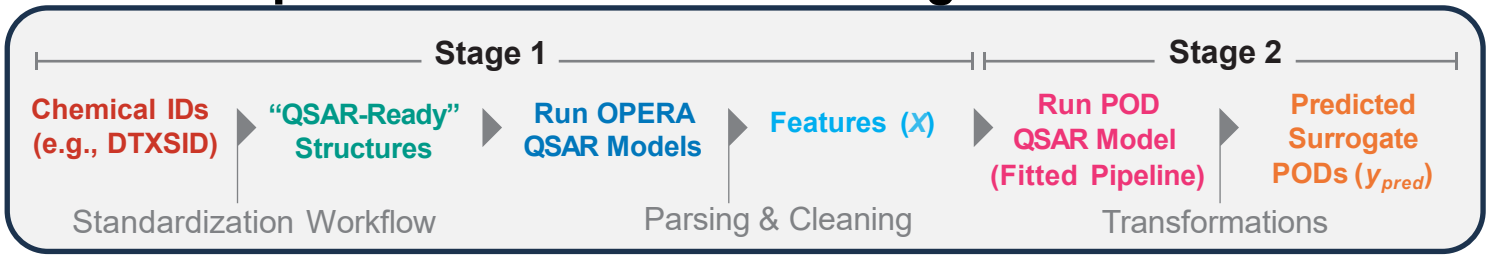


Why a two-stage model?

- 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})

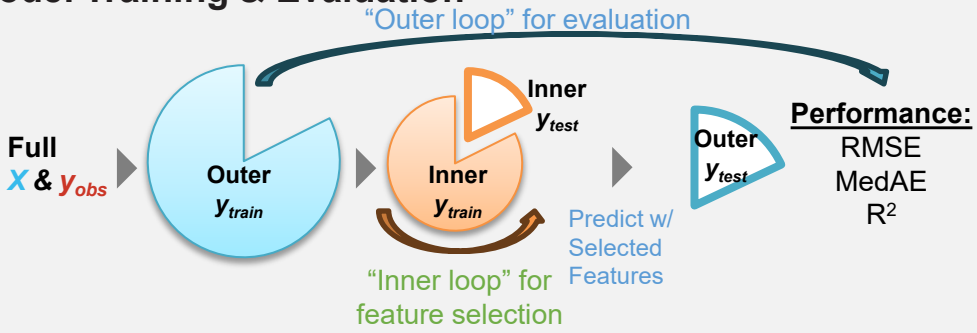
n_g = 5,209
n_{rd} = 4,938

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

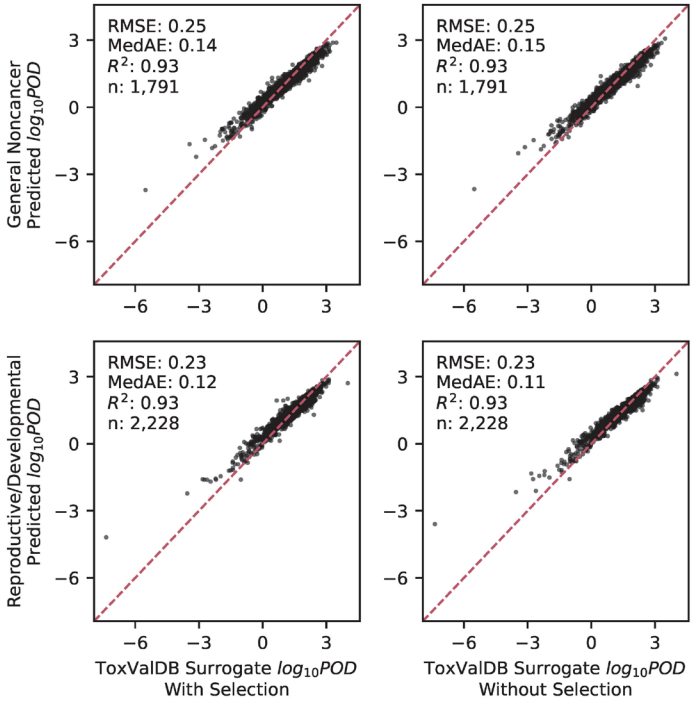


Performance:
RMSE
MedAE
R^2

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

Kvasnicka et al. 2024

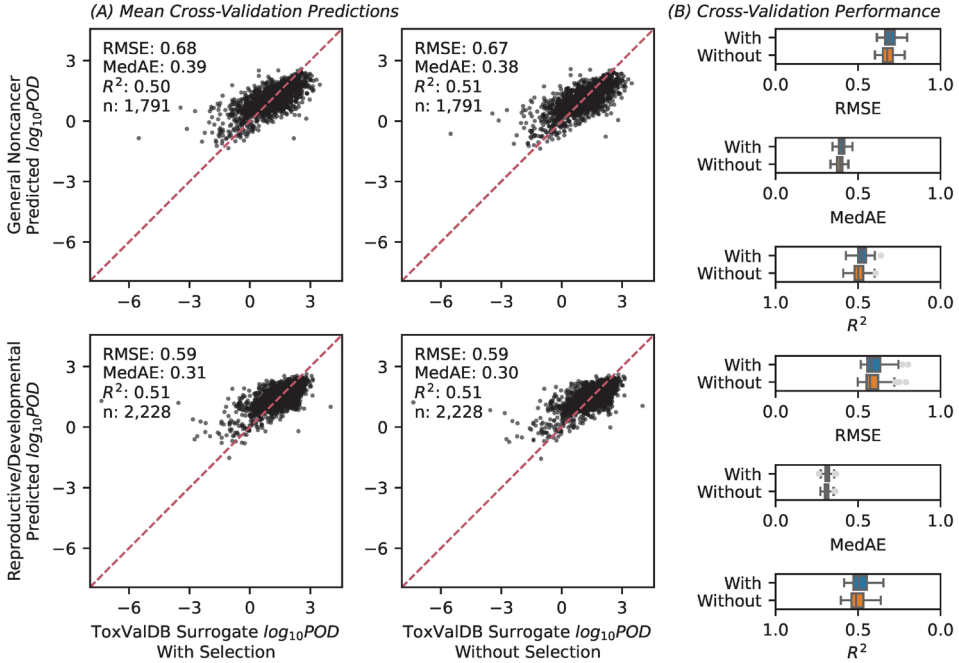
In-Sample Model Fit: High Accuracy for Training Chemicals



- **RMSE (Root Mean Squared Error)**: Measures the average prediction error magnitude. Lower values indicate better accuracy.
- **MedAE (Median Absolute Error)**: Robust to outliers. Lower values indicate better accuracy.
- **R^2 (Coefficient of Determination):** Indicates how well the model explains variance in the data. Values closer to 1 show better fit.

Kvasnicka et al. 2024

Good Out-of-Sample (Cross-Validation) Performance



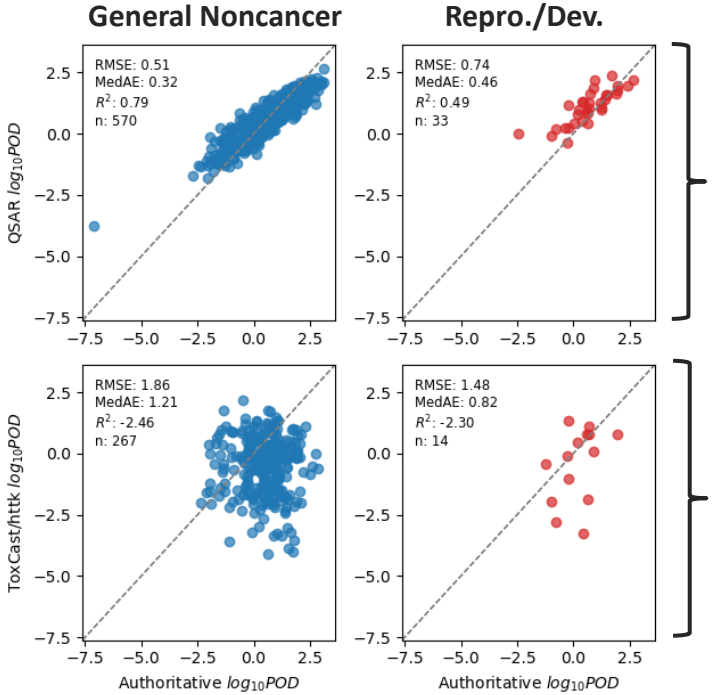
Expected Performance

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

[Kvasnicka et al. 2024](#)

VETERINARY MEDICINE & BIOMEDICAL SCIENCES
TEXAS A&M UNIVERSITY

QSAR PODs Correlated Better with Regulatory PODs than ToxCast & *In Vitro* NAMs



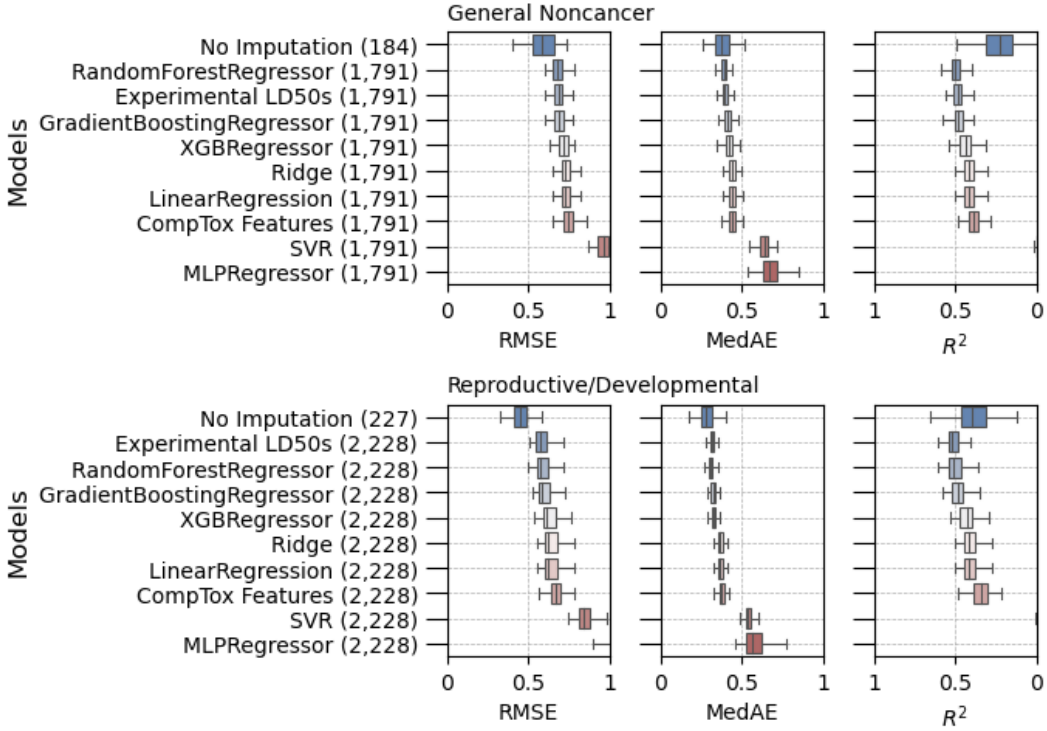
QSAR: Good correspondence with regulatory PODs

ToxCast/httk: Performed worse than a naïve constant model

Sensitivity Analysis: Random Forest with OPERA Features Outperformed Other Models

Models Compared

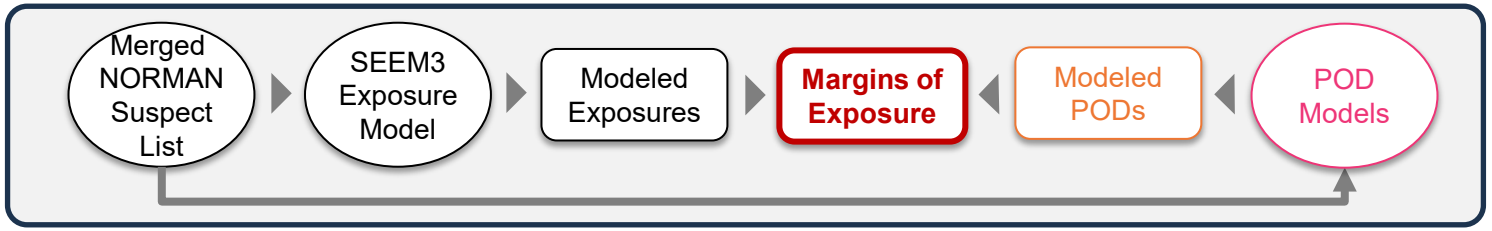
1. No feature selection applied
2. Alternative machine learning estimators for model fitting
3. Alternative features for modeling
   - Experimental LD50s instead of predicted
   - Features from OPERA, TEST on EPA CompTox, and RDKit 2D descriptors
   - No imputation of missing values



Kvasnicka et al. 2024

Model Application: Derived Margins of Exposure for ~30,000 Environmental Chemicals for Risk Screening

Compared Predicted *PODs* with Predicted *Exposures*

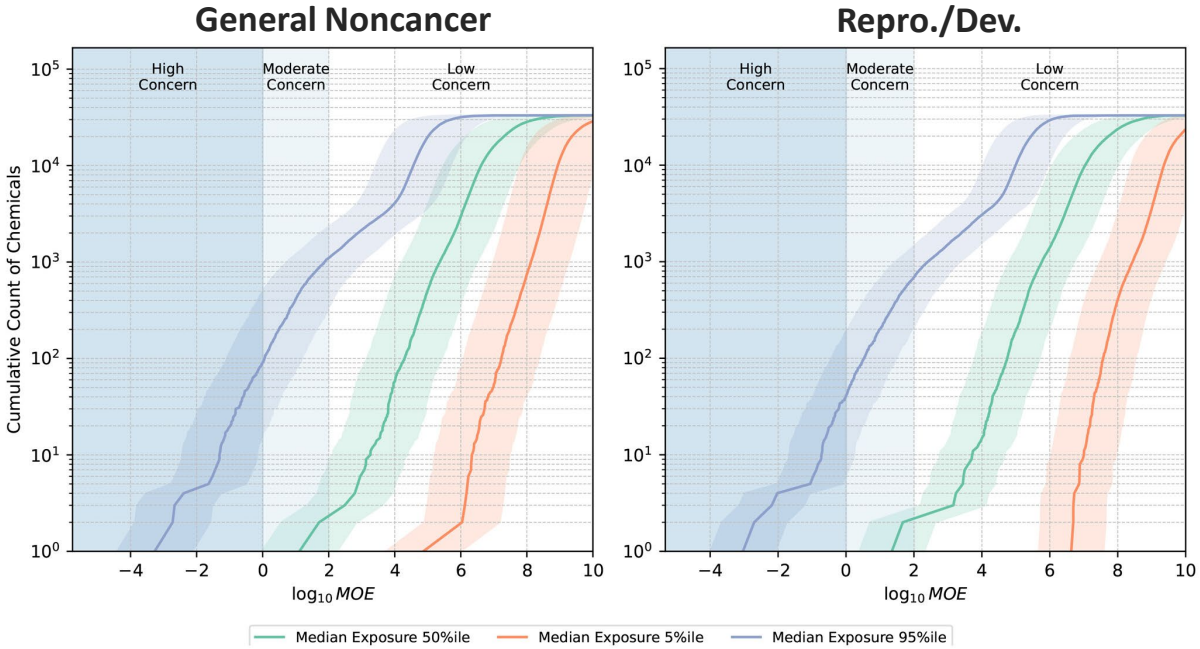


SEEM3: Ring et al. 2018

$$\text{Margin of Exposure (MOE)} = \frac{\text{POD}}{\text{Exposure}}$$

where *lower* MOE indicates *higher* risk

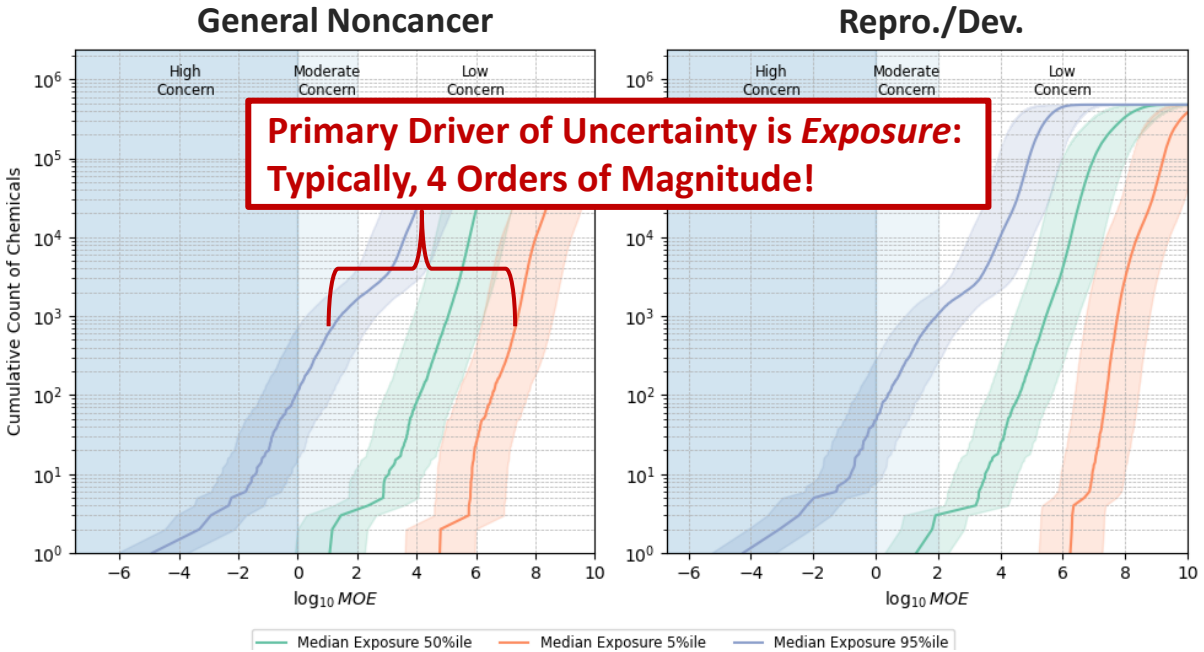
MOEs Revealed Several Thousand Chemicals of Concern: Should Prioritize Further Investigation



Kvasnicka et al. 2024

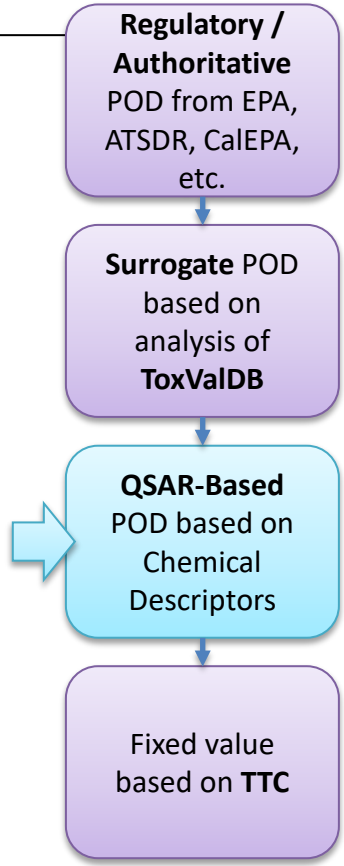
New Web App Makes Predictions Available: 800K+ Chemicals from EPA CompTox Dashboard

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



Conclusion

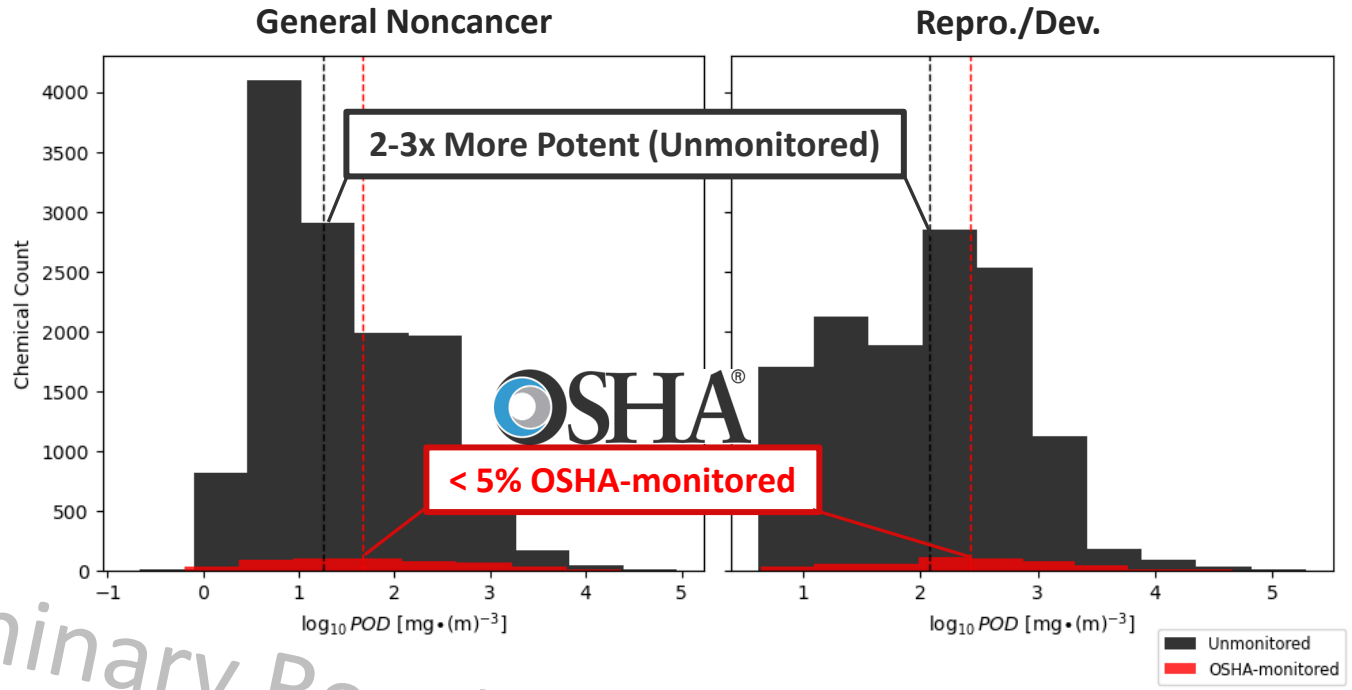
- PODs are absent for tens-hundreds of thousands of chemicals for characterizing human health risks and impacts
- Regulatory/authoritative PODs cover a very limited set of chemicals
- Machine learning can substantially expand the coverage of chemicals with actionable PODs
- PODs for *inhalation* are still unknown, and exposure appears to be the primary uncertainty.
- **Current Work:** Applying the approach to inhalation, and to better model inhalation exposure.



Current Work at EPA: How to Protect Workers from Thousands of Potential Chemicals?

- Toxic Substances Control Act / TSCA (1976)
  - Regulation of chemicals in commerce
- Lautenberg Act (2016)
  - Protect highly-exposed subpopulations
  - Including workers, esp. inhalation
- ~29,000 commercially active chemicals
  - **Challenge for exposure monitoring**

Applying Inhalation POD Models to TSCA Chemicals Highlights a Gap in Exposure Monitoring



Preliminary Results

Conclusion

- PODs are absent for tens-hundreds of thousands of chemicals for characterizing human health risks and impacts
- Regulatory/authoritative PODs cover a very limited set of chemicals
- Machine learning can substantially expand the coverage of chemicals with actionable PODs
- PODs for *inhalation* are still unknown, and exposure appears to be the primary uncertainty.
- **Current Work:** Applying the approach to inhalation, and to better model inhalation exposure.