New Approach Methodologies for Exposure from EPA’s ExpoCast Project

Kristin Isaacs

National Exposure Research Laboratory
Office of Research and Development, United States Environmental Protection Agency

SCIENTIFIC ADVISORY COMMITTEE ON ALTERNATIVE TOXICOLOGICAL METHODS
Arlington, VA
September 19, 2019
Disclaimer

The views expressed in this presentation are those of the authors and do not necessarily reflect the views or policies of the U.S. EPA
EPA is charged with evaluating risks associated with 1000s of chemicals in commerce

- For example, as of June 2017 there were 67,709 chemicals on the TSCA Inventory

Evaluating chemicals for risk to humans or the environment requires information on hazard and exposure potential

- Exposure potential quantifies the degree of contact between a chemical and a receptor

Toxicokinetic information is required to bridge hazard and exposure (what real-world exposure is required to produce an internal concentration consistent with a potential hazard?)
Risk is Multifaceted

EPA's ExpoCast Project

Hazard

Toxicokinetics

Exposure

Potential Hazard from in vitro with Reverse Toxicokinetics

Potential Exposure

mg/kg BW/day

Lower Risk  Medium Risk  Higher Risk
Forecasting Exposure is a Systems Problem

- **USE and RELEASE**
  - Consumer Products and Durable Goods
  - Chemical Manufacturing and Processing
- **MEDIA**
  - Indoor Air, Dust, Surfaces
  - Waste
  - Food
  - Drinking Water
  - Outdoor Air, Soil, Surface and Ground Water
- **RECEPTOR**
  - Human
  - Ecological Flora and Fauna
Forecasting Exposure is a Systems Problem

EPA Considers 4 Pathway Types:
- Consumer
- Occupational
- Ambient
- Ecological

USE and RELEASE

MEDIA

RECEPTOR

Consumer Products and Durable Goods

Other Industry

Chemical Manufacturing and Processing

Industrial Releases

Waste

Food

Drinking Water

Outdoor Air, Soil, Surface and Ground Water

Ecological Flora and Fauna

Human

Indoor Air, Dust, Surfaces

Direct Use (e.g., surface cleaner)

Occupational Use
Forecasting Exposure is a Systems Problem

Forward Models for Pathways

USE and RELEASE

MEDIA

EXPOSURE (MEDIA + RECEPTOR)

CONSUMER PRODUCTS AND DURABLE GOODS

OTHER INDUSTRY

CHEMICAL MANUFACTURING AND PROCESSING

Indoor Air, Dust, Surfaces

Near-Field Direct

Near-Field Indirect

Occupational

Dietary

Far-Field

Ecological

Human

Ecological Flora and Fauna

Exposure (e.g., chemical intake in mg/kg/day) is unobservable
Forecasting Exposure is a Systems Problem

Forward Models for Pathways

USE and RELEASE

MEDIA

EXPOSURE (MEDIA + RECEPTOR)

RECEPTOR

MONITORING DATA

Evaluation

Exposure (e.g., chemical intake in mg/kg/day) is unobservable
Traditional Exposure Data Are Limited

The ExpoCast project and its collaborators are working to fill gaps in exposure data for 1000s of chemicals using high-throughput new approach methodologies (NAMs) for exposure.

Use category information for chemicals being tested via high throughput screening at EPA.

*Egeghy et al. (2012)*
New Approach Methodologies and Exposure

“NAMs were taken in a broad context to include in silico approaches, in chemico and in vitro assays, as well as the inclusion of information from the exposure of chemicals in the context of hazard assessment.”

“...the committee sees the potential for the application of computational exposure science to be highly valuable and credible for comparison and priority-setting among chemicals in a risk-based context.”
Eight Classes of NAMs for Exposure

- **Chemical descriptors** that provide information on chemicals in an exposure context (e.g., how chemicals are used)
- **Machine-learning approaches** that use these descriptors to fill gaps in existing data
- **High-throughput exposure models** for various pathways
- **High-throughput measurements** to fill gaps in monitoring data
- High-throughput approaches for measuring and predicting chemical **toxicokinetics**
- New **evaluation frameworks** for integrating models and monitoring to provide consensus exposure predictions
- All these pieces together provide the tools for high-throughput **chemical prioritization**
Chemical Descriptor NAMs

CPCat

General use categories

CPCPdb

Reported chemicals in products

Ingredient Lists

Chemical role in products

Functional Use Data

Measured Data

Identification of Compounds

Chemical role in products

https://comptox.epa.gov/dashboard
Chemical Descriptor NAMs

- We are using informatics approaches to obtain and curate additional chemical descriptor information
- Data from chemical use, monitoring, and release domains
- Public data sources: reports, open literature, databases
- Focus on expanding to critical pathways (e.g., occupational)
High-throughput Measurement NAMs

- **Targeted Analysis:**
  - We know exactly what we’re looking for
  - 10s – 100s of chemicals

- **Non-Targeted Analysis (NTA):**
  - We have no preconceived lists
  - 1,000s – 10,000s of chemicals

- **Ongoing consumer product scanning and blood sample monitoring via contract (NTA and confirmation of tentative IDs with available standards**

- **Development of significant in-house capabilities**

- **EPA is coordinating a comparison of non-targeted screening workflows used by leading academic and government groups using known chemical mixtures (ToxCast) and standardized environmental/biological samples**

- **Goal is to develop tools, databases, and workflows for rapid analysis of any sample for chemicals of interest, i.e. exposure forensics**
NTA Studies in the ExpoCast Project

**Consumer Products**
- Pilot study of 20 classes of products, including clothing, personal care products, carpet padding, cleaners
- 5 products from each category
- GC X GC/MS
- Phillips et al., *Env. Sci. Tech.* 2018

**Recycled Materials**
- Products from six categories of recycled products, including building materials, paper products, toys, and clothing
- 20 products per category with 50% recycled material, 8 products virgin materials
- GC X GC/MS
- Lowe et al., in prep

**Residential Dust**
- Dust samples from 56 U.S. homes in the American Health Homes Survey
- LC/MS
- Rager et al., *Env. Int.*, 2016
Machine Learning Based Classification Models
(Random Forest, Breiman, 2001)

Prediction of Potential Alternatives from Chemical Libraries

Phillips et al. (2017)
High Throughput Exposure Model NAMs

Consumer
Isaacs et al. (2014)

Occupational
Meyer et al. (2017)

Ambient

Ecological
Biryol et al. (2017)

Barber et al. (2017)
Evaluation NAMs: The SEEM Framework

- We use Bayesian methods to incorporate multiple pathway models into consensus predictions for 1000s of chemicals within the **Systematic Empirical Evaluation of Models (SEEM)**

---

Evaluate Model Performance and Refine Models

Evaluate Inference

Inferred Intake Rate

Estimate Uncertainty

Calibrate models

Different Chemicals

Available Exposure Predictors

Exposure

Chemicals with Monitoring Data

Dataset 1

Dataset 2

Model 1

Model 2

Apply calibration and estimated uncertainty to other chemicals

Space of Chemicals

Hurricane path prediction is an example of integrating multiple models

Material from John Wambaugh
Evaluation NAMs: The SEEM Framework

- We use Bayesian methods to incorporate multiple pathway models into consensus predictions for 1000s of chemicals within the **Systematic Empirical Evaluation of Models (SEEM)**

- ExpoCast is developing SEEM Models for multiple receptors and pathways
- **Human** (Wambaugh et al., 2013, 2014; Ring et al., 2018)
  - Evaluated with NHANES biomonitoring data
- **Ecological** (Sayre et al., *in prep*)
  - Evaluated with USGS water data
- **Occupational** (planned)
  - Evaluated with OSHA occupational monitoring data
<table>
<thead>
<tr>
<th>Predictor (including Models)</th>
<th>Reference(s)</th>
<th>Chemicals Predicted</th>
<th>Pathways</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPA Inventory Update Reporting and Chemical Data Reporting (CDR) (2015)</td>
<td>US EPA (2018)</td>
<td>7856</td>
<td>All</td>
</tr>
<tr>
<td>United Nations Environment Program and Society for Environmental Toxicology and Chemistry toxicity model (USEtox) Industrial Scenario (2.0)</td>
<td>Rosenbaum et al. (2008)</td>
<td>8167</td>
<td>Far-Field Industrial</td>
</tr>
<tr>
<td>USEtox Pesticide Scenario (2.0)</td>
<td>Fantke et al. (2011, 2012, 2016)</td>
<td>940</td>
<td>Far-Field Pesticide</td>
</tr>
<tr>
<td>Risk Assessment IDentification And Ranking (RAIDAR) Far-Field (2.02)</td>
<td>Arnot et al. (2008)</td>
<td>8167</td>
<td>Far-Field Pesticide</td>
</tr>
<tr>
<td>Fugacity-based INdoor Exposure (FINE) (2017)</td>
<td>Bennett et al. (2004), Shin et al. (2012)</td>
<td>645</td>
<td>Residential</td>
</tr>
<tr>
<td>RAIDAR-ICE Near-Field (0.803)</td>
<td>Arnot et al., (2014), Zhang et al. (2014)</td>
<td>1221</td>
<td>Residential</td>
</tr>
<tr>
<td>USEtox Residential Scenario (2.0)</td>
<td>Jolliet et al. (2015), Huang et al. (2016,2017)</td>
<td>615</td>
<td>Residential</td>
</tr>
<tr>
<td>USEtox Dietary Scenario (2.0)</td>
<td>Jolliet et al. (2015), Huang et al. (2016), Ernstoff et al. (2017)</td>
<td>8167</td>
<td>Dietary</td>
</tr>
</tbody>
</table>

Material from John Wambaugh
- Machine learning models were built for each of four exposure pathways
- Pathway predictions can be used for large chemical libraries
- Use prediction (and accuracy of prediction) as a prior for Bayesian analysis
- Each chemical may have exposure by multiple pathways
We extrapolate to predict relevant pathway(s), median intake rate, and credible interval for each of 479,926 chemicals.

Of the chemicals evaluated, 30% have less than a 50% probability for relevance to any of the four pathways and are considered outside the “domain of applicability.”

This approach identifies 1,880 chemicals for which the median population intake rates may exceed 0.1 mg/kg bodyweight/day.
• We extrapolate to predict relevant pathway(s), median intake rate, and credible interval for each of 479,926 chemicals.

• Of the chemicals evaluated, 30% have less than a 50% probability for exposure via any of the four pathways and are considered outside the “domain of applicability.”

• This approach identifies 1,880 chemicals for which the median population intake rates may exceed 0.1 mg/kg bodyweight/day.

• There is 95% confidence that the median intake rate is below 1 µg/kg BW/day for 474,572 compounds.
Toxicokinetics NAMs

- Chemical-specific data for toxicokinetics (TK) are as sparse as for exposure
- High throughput TK methods have provided data for nearly 1000 chemicals over the past decade
- However, thousands of chemicals remain requiring machine learning and QSAR approaches

Rotroff et al. (2010) 35 chemicals
Wetmore et al. (2012) +204 chemicals
Wetmore et al. (2015) +163 chemicals
Wambaugh et al. (in prep.) + ~300 chemicals

**In vitro Measurements**

Generic Physiologically-Based Toxicokinetic Models

Machine Learning Models

Material from John Wambaugh
High throughput *in vitro* screening + toxicokinetics NAMs can estimate doses needed to cause bioactivity (e.g., Wetmore *et al.*, 2015)

Consensus exposure rates with uncertainty (e.g., Ring *et al.*, 2018)
Summmary

• Estimates of human and ecological exposures are required as critical input to risk-based prioritization and screening of chemicals.

• The ExpoCast project seeks to develop the data, tools, and evaluation approaches required to generate rapid and scientifically-defensible:
  – Exposure predictions for the full universe of existing and proposed commercial chemicals.
  – The toxicokinetic data required to relate bioactive concentrations identified in high-throughput screening to predicted real world doses (i.e. in vitro-in vivo extrapolation).

• We are developing and applying computational and analytical new approach methodologies for exposure science and toxicokinetics that are appropriate for application to 1000s of chemicals.

• Rapid prediction of chemical exposure and bioactive doses allows prioritization based upon risk.
ExpoCast Project
(Exposure Forecasting)


