Updates on EPA’s High-Throughput Exposure Forecast (ExpoCast) Research project,

John Wambaugh
National Center for Computational Toxicology
U.S. EPA, Office of Research and Development

CompTox Community of Practice, November 20, 2014

The views expressed in this presentation are those of the author and do not necessarily reflect the views or policies of the U.S. EPA
Introduction

- The timely characterization of the human and ecological risk posed by thousands of existing and emerging commercial chemicals is a critical challenge facing EPA in its mission to protect public health and the environment.

- ExpoCast is an EPA ORD initiative to develop the necessary approaches and tools for rapidly predicting exposure for thousands of chemicals (Cohen-Hubal, et al., 2010).

- **Proof of Concept (First Generation Analysis):** Used off-the-shelf high throughput exposure models – simple description of near field exposure predicted more than existing HT models (Wambaugh et al., 2013).

“All cases are unique, and very similar to others.” — T.S. Eliot
Risk-based Prioritization Requires Exposure

- **Tox21/ToxCast**: Examining thousands of chemicals using high throughput screening assays to identify *in vitro* concentrations that perturb biological pathways (Schmidt, 2009)

- In Wetmore *et al.* (2012), High throughput toxicokinetic *in vitro* methods are used to approximately convert *in vitro* bioactive concentrations (µM) into daily doses needed to produce similar levels in a human (mg/kg BW/day)

- These doses can then be directly compared with exposure rates, *where available*
Studies like Wetmore et al. (2012) addressed the need for toxicokinetic data.
As in Egeghy et al. (2012), there is a paucity of data for providing context to HTS data.
Exposure Science in the 21st Century

- 2012 NRC report:
  - New tools needed for screening and prioritization of chemicals for targeted toxicity testing
  - New, focused exposure assessments or monitoring studies needed
  - Better quantification of population vulnerability needed

Figure from Egeghy et al. (2012), “The exposure data landscape for manufactured chemicals”
Exposure Space

Chemical Manufacture

Environment al Release

Consumer Products, Articles, Building Materials

Direct Use (e.g., lotion)

Residential Use (e.g., flooring)

Air, Dust, Surfaces

Waste

Food

Air, Soil, Water

Ecological Flora and Fauna

Human

Biomarkers of Exposure

Media Samples

Biomarkers of Exposure

Figure from Kristin Isaacs
Exposure Pathways

- Chemical Manufacture
- Environment al Release
- Consumer Products, Articles, Building Materials
- Near-Field Direct
- Near-Field Indirect
- Residential Use (e.g., flooring)
- Direct Use (e.g., lotion)
- Waste
- Air, Dust, Surfaces
- Food
- Air, Soil, Water
- Media Samples
- Human
- Biomarkers of Exposure
- Ecological Flora and Fauna
- Biomarkers of Exposure
- Monitoring Data
- Receptors

Figure from Kristin Isaacs
Forward Modeling of Exposure Pathways

Data and Models

EXPOSURE PATHWAY (MEDIA + RECEPTOR)

Near-Field Direct
Near-Field Indirect

RECEPTORS

MONITORING DATA

Human

Biomarkers of Exposure

Media Samples

Biomarkers of Exposure

Human

Chemical Manufacture

Consumer Products, Articles, Building Materials

Environmental Release

Air, Soil, Water

Food

Air, Dust, Surfaces

Waste

Direct Use (e.g., lotion)

Residential Use (e.g., flooring)

Dietary

Far-Field

Ecological Flora and Fauna

Ecological

Figure from Kristin Isaacs
Forward Predicting Exposure

Components of the source-to-outcome continuum:

- **Exposure estimate**
- **Dose estimate**
- **BR dose estimate**

Exposed source → Environmental measurements → Statistical models → Biomarker measurements → Statistical models → BR biomarker measurements → Health outcome

Exposure research → Biomonitoring research → Health effects research

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Key</th>
<th>Parameter</th>
<th>Definition</th>
</tr>
</thead>
</table>
| △      | Estimated Value | 1) Exposure estimate  
2) Dose estimate  
3) BR dose estimate | 1) Estimated mass of a chemical that comes into contact with a human over time  
2) Estimated mass of a chemical inside a human over time  
3) Estimated amount of the dose at a specific target inside a human |
| □      | Measured value | 1) Environmental measurement  
2) Biomarker measurement  
3) BR biomarker measurement | 1) Observation of a stressor in environmental media that reflects a source  
2) Observation of a stressor in biological media that reflects an exposure/dose  
3) Observation of a stressor in biological media that reflects a BR dose |
|        | Empirical model | 1) Statistical model (blue)  
2) Exposure model (red)  
3) Kinetic model (red)  
4) Dynamic model (red) | 1) Model that evaluates observed variables for hypothesis testing  
2) Model that estimates exposure using environmental measurements and exposure factors  
3) Model that describes how stressors enter and is removed from a human  
4) Model that describes the effect of stressor on the human body |

Sobus et al. (2011)
Inferring Exposure

Tan et al. (2012)
Evaluation of Forward Predictions with Inferred Exposure

Data and Models

- Chemical Manufacture
- Environment al Release
- Air, Soil, Water
- Food
- Waste

EXPOSURE PATHWAY
(MEDIA + RECEPTOR)

- Near-Field Direct
- Near-Field Indirect
- Human
- Biomarkers of Exposure
- Media Samples
- Biomarkers of Exposure
- Ecological Flora and Fauna

Data and Models

- MONITORING DATA
- RECEPTORS
- Food
- Near-Field Direct
- Residential Use (e.g., flooring)
- Waste

Figure from Kristin Isaacs
Investigating Exposure to Environmental Chemicals

Tan et al. (2012):

A cartoon illustrating the relation of different factors and knowledge domains in the exposure reconstruction process. This cartoon is generated using key terms in this review and their semantic/lexical relationships using the visual analysis of IBM’s www.many-eyes.com Phrasenet analysis.
Sobus et al. (2011): Use a mix of empirical and mechanistic models

Empirical models can be as simple as “rule of thumb”, i.e. heuristics of exposure
How to Make Good Forecasts

1) Think probabilistically (especially, Bayesian): We use an approach that evaluates model performance systematically across as many chemicals (and chemistries) as possible.

2) Forecasts change: Today’s forecast reflects the best available data today but we must accept that new data and new models will cause predictions to be revised.

3) Look for consensus: We evaluate as many models and predictors/predictions as possible.

Orrin Pilkey & Olinda Pilkey-Jarvis (2007)

Nate Silver (2012)
Exposure Forecasting (ExpoCast)

• Develop the tools and data necessary to rapidly quantify human and ecological exposure potential of chemicals

• Focus is distinct from many existing exposure tools that support either screening level assessments on a per chemical basis or full regulatory risk assessment

In Nate Silver’s terminology:

a prediction is a specific statement
a forecast is a probabilistic statement

Wikipedia (statistics): “when information is transferred across time, often to specific points in time, the process is known as forecasting”
Systematic Empirical Evaluation of Models (SEEM)

• There are four basic steps in the SEEM framework
  1. Forward prediction of exposures, which involves model curation and parameterization
  2. Inference of exposures from monitoring data
  3. Systematic evaluation and calibration of the predictions against the inferred exposures
  4. Extrapolation of the calibrated model predictions and estimated uncertainty to chemicals with no monitoring data.

• To achieve these aims the SEEM framework used Bayesian formalism and multivariate, linear regression for demonstrating and evaluating predictive ability
Illustration of the SEEM Framework

Apply calibration and uncertainty to other chemicals

EDSP Chemicals
QSARs and HTE Data

Biomonitoring Data
Dataset 1
Dataset 2

Exposure Inference

Inferred (Reverse) Exposure

Estimate Uncertainty

Calibrate models

Forward Predictions

Evaluate Model Performance and Refine Models

Model 1
Model 2

...
Goals for High Throughput Exposure

- Incorporate multiple models into consensus predictions for 1000s of chemicals
- Evaluate/calibrate predictions with available measurement data across many chemical classes
- Empirically estimate uncertainty in predictions
Data Availability for Evaluating Model Predictions

- Currently we use the CDC NHANES urine data
- Many chemicals had median conc. below the limit of detection (LoD)
  - Most chemicals >LoD not high production volume
- 106 chemicals inferred from urine to date
- Dozens more expected with serum/blood model

Wambaugh et al. (2013)
Applying ExpoCast

- There are 1000s of chemicals to which we might be exposed
- How can we use ExpoCast to pick chemicals with more likely exposure?
- What about uncertainty?
Applying ExpoCast

NHANES

- The CDC targets some chemicals for exposure biomonitoring
Applying ExpoCast

- They find evidence of high exposures for some chemicals
- Moderate exposures for others
- And many chemicals are below the limit of detection
Applying ExpoCast

- We use the chemical descriptors and high level use information (ACToR UseDB) that is available for thousands of EDSP chemicals to organize the NHANES chemicals.

Images from Thinkstock
Applying ExpoCast

- We can then predict which chemicals without monitoring data are most like high, moderate, and low exposure NHANES chemicals.

There will still be other chemicals without characteristics that are predictive of NHANES chemicals.
NHANES is Much More than a Chemical Survey

• Separate evaluations can be done for various demographics

<table>
<thead>
<tr>
<th>Age group</th>
<th>Survey years</th>
<th>Geometric mean (95% conf. interval)</th>
<th>50th</th>
<th>75th</th>
<th>90th</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-11 years</td>
<td>03-04</td>
<td>3.55 (2.95-4.29)</td>
<td>3.80 (2.70-5.00)</td>
<td>6.90 (6.00-8.30)</td>
<td>12.6 (9.50-16.1)</td>
</tr>
<tr>
<td></td>
<td>05-06</td>
<td>2.86 (2.52-3.24)</td>
<td>2.70 (2.30-2.90)</td>
<td>5.00 (4.40-5.80)</td>
<td>13.5 (9.30-16.8)</td>
</tr>
<tr>
<td></td>
<td>07-08</td>
<td>2.46 (2.20-2.75)</td>
<td>2.40 (1.90-3.00)</td>
<td>4.50 (3.70-5.50)</td>
<td>7.00 (6.30-8.30)</td>
</tr>
<tr>
<td>12-19 years</td>
<td>03-04</td>
<td>3.74 (3.31-4.22)</td>
<td>4.30 (3.60-4.80)</td>
<td>7.80 (6.50-9.90)</td>
<td>13.5 (11.8-16.8)</td>
</tr>
<tr>
<td></td>
<td>05-06</td>
<td>2.42 (2.18-2.68)</td>
<td>2.40 (2.10-2.70)</td>
<td>4.30 (3.90-5.20)</td>
<td>8.40 (6.50-8.80)</td>
</tr>
<tr>
<td></td>
<td>07-08</td>
<td>2.44 (2.14-2.78)</td>
<td>2.30 (2.10-2.60)</td>
<td>4.40 (3.70-5.70)</td>
<td>9.70 (7.30-11.8)</td>
</tr>
<tr>
<td>20 years and older</td>
<td>03-04</td>
<td>2.41 (2.15-2.72)</td>
<td>2.60 (2.30-2.80)</td>
<td>5.10 (4.50-5.70)</td>
<td>9.50 (8.10-10.1)</td>
</tr>
<tr>
<td></td>
<td>05-06</td>
<td>1.75 (1.62-1.89)</td>
<td>1.80 (1.70-2.00)</td>
<td>3.40 (3.10-3.70)</td>
<td>6.40 (5.80-7.30)</td>
</tr>
<tr>
<td></td>
<td>07-08</td>
<td>1.99 (1.82-2.18)</td>
<td>2.00 (1.80-2.30)</td>
<td>3.90 (3.40-4.60)</td>
<td>7.40 (6.60-8.60)</td>
</tr>
</tbody>
</table>

CDC, Fourth National Exposure Report (2011)
Linking NHANES Urine Data and Exposure

**Generalization of LaKind and Naiman (2008)**

Parent

Products

1

2

3

4

$$\text{Steady-state assumption}$$

$$(\text{mg/kg/day})_i = \frac{1}{70 \text{ kg}} \frac{\text{mg}_i}{\text{g}_{\text{creatin}}} \times \frac{\text{g}_{\text{creatin}}}{\text{day}}$$

$$(\text{mg/kg/day})_0 = MW_0 \sum_i \phi_{oi} \frac{(\text{mg/kg/day})_i}{MW_i}$$

Method detailed in Wambaugh et al. (2013) Supplemental
Linking NHANES Urine Data and Exposure

Generalization of LaKind and Naiman (2008)

Parent → Products

1

2

3

4

Steady-state assumption

\[
\left( \frac{\text{mg/kg/day}}{} \right)_i = \frac{1}{70 \text{ kg}} \frac{\text{mg}_i}{\text{g}_{\text{creatinine}}} \times \frac{\text{g}_{\text{creatinine}}}{\text{day}}
\]

Observations (CDC NHANES urine samples)

\[
\left( \frac{\text{mg/kg/day}}{} \right)_0 = M W_0 \sum_i \phi_{0i} \left( \frac{\text{mg/kg/day}}{} \right)_i \frac{\text{MW}_i}{\text{MW}_0}
\]

Method detailed in Wambaugh et al. (2013) Supplemental
Linking NHANES Urine Data and Exposure

Generalization of LaKind and Naiman (2008)

Parent

Products

1

2

3

4

Steady-state assumption

\[
\left( \frac{\text{mg/kg/day}}{} \right)_i = \frac{1}{70 \text{ kg}} \frac{\text{mg}_i}{\text{g}_{\text{creatinne}}} * \frac{\text{g}_{\text{creatinne}}}{\text{day}}
\]

\[
\left( \frac{\text{mg/kg/day}}{} \right)_0 = M W_0 \sum_i \phi_{0i} \left( \frac{\text{mg/kg/day}}{} \right)_i \frac{M W_i}{M W_i}
\]

Unknwonns (we choose to use Bayesian analysis via Markov Chain Monte Carlo or MCMC)

Observations (CDC NHANES urine samples)

Method detailed in Wambaugh et al. (2013) Supplemental
Lots of unknowns, but MCMC can try many possibilities (JAGS and STAN samplers used)

Method detailed in Wambaugh et al. (2013) Supplemental
Limit of Detection (LOD)

- If observations < analytic detection limits: We model the data as left censored observations from lognormal population distribution.

Parameters for distribution: log geometric mean (\(\ln(GM)\)) and standard deviation.
- We also estimate these parameters with MCMC.
- Generally, these estimates have greater uncertainty.

Figure from Woodrow Setzer
Systematic Empirical Evaluation of Models

Apply calibration and uncertainty to other chemicals

EDSP Chemicals
- QSARs and HTE Data

Biomonitoring Data
- Dataset 1
- Dataset 2

Exposure Inference

Dataset 1
Dataset 2

Model 1
Model 2

Inferred (Reverse) Exposure

Forward Predictions

Evaluate Model Performance and Refine Models

Calibrate models

Estimate Uncertainty
Statement of New Problem: Data Concerns

• If a simple near-field/far-field heuristic was most predictive so far (Wambaugh et al, 2013), then do there exist other heuristics with the power to distinguish chemicals with respect to exposure?

• What we would like to know is:

  • What are the few, most-easily obtained exposure heuristics that allow for prioritization?
Statement of New Problem: Data Concerns

• If a simple near-field/far-field heuristic was most predictive so far (Wambaugh et al, 2013), then do there exist other heuristics with the power to distinguish chemicals with respect to exposure?

• What we would like to know is:
  • What are the few, most-easily obtained exposure heuristics that allow for prioritization?

• What we can answer is this:
  • Given a variety of rapidly obtained data (putative use categories and physico-chemical properties, largely from QSAR) which data best explain exposure inferred from the available biomonitoring data?
  • Hoping to find simple heuristics for exposure e.g., use in fragrances, use as a food additive, octanol:water partition coefficient, vapor pressure
ACToR UseDB: Chemical Use Categories estimated from ACToR (computational toxicology database):

- The sources for chemical data were assigned to various chemical use categories.
- Chemicals from multiple sources were assigned to multiple categories.

**Table: Hits per use category for a given chemical**

<table>
<thead>
<tr>
<th>CASRN</th>
<th>Category 1</th>
<th>Category 2</th>
<th>...</th>
<th>Category 12</th>
</tr>
</thead>
<tbody>
<tr>
<td>65277-42-1</td>
<td>0</td>
<td>10</td>
<td>...</td>
<td>1</td>
</tr>
<tr>
<td>50-41-9</td>
<td>31</td>
<td>7</td>
<td>...</td>
<td>3</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

**12 Chemical Use Categories**

- Antimicrobials
- Chemical Industrial Process
- Consumer
- Dyes and Colorants
- Fertilizers
- Food Additive
- Fragrances
- Herbicides
- Personal Care Products
- Pesticides
- Petrochemicals
- Other

http://actor.epa.gov/cpcat/

Work by Alicia Frame, Kathie Dionisio, Richard Judson
Dionisio et al. manuscript in preparation
High Throughput Descriptors for Exposure

Noisy data and the danger of over-fitting

- The average relative AIC (smaller is better) for models made with different numbers of parameters for explaining 1500 different combinations of chemical exposures

Yes / No Use Descriptors

Physico-chemical Properties (EPI Suite)

Wambaugh et al., (2014)
Not All Descriptors Are Useful

- The average relative AIC (smaller is better) for models made with different numbers of parameters for explaining 1500 different combinations of chemical exposures.

- The predictors involved in the optimal model with higher frequencies are represented by darker circles, and those with lower frequencies by lighter circles.

- As a sanity check, two random variables generated from binomial distribution with probability 50% and 10% of obtaining 1, are not selected as optimal descriptors in the five factor model.

Wambaugh et al., (2014)
**Systematic Empirical Evaluation of Models**

1. **EDSP Chemicals**
   - QSARs and HTE Data

2. **Biomonitoring Data**
   - Dataset 1
   - Dataset 2

3. **Exposure Inference**

4. **Inferred (Reverse) Exposure**

5. **Evaluate Model Performance and Refine Models**

6. **Calibrate models**

Apply calibration and uncertainty to other chemicals

- **Forward Predictions**
- **Evaluate Model Performance and Refine Models**

- **Estimate Uncertainty**

![Diagram showing the systematics of empirical evaluation of models, including data sources, exposure inference, and model calibration and evaluation.](image-url)
R² ≈ 0.5 indicates that we can predict 50% of the chemical to chemical variability in mean NHANES exposure rates.

Same five predictors work for all NHANES demographic groups analyzed – stratified by age, sex, and body-mass index.
# High-throughput exposure heuristics

<table>
<thead>
<tr>
<th>Heuristic</th>
<th>Description</th>
<th>Number of Chemicals</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACToR “Consumer use &amp; Chemical/Industrial Process use”</td>
<td>Chemical substances in consumer products (e.g., toys, personal care products, clothes, furniture, and home-care products) that are also used in industrial manufacturing processes. Does not include food or pharmaceuticals.</td>
<td>37   683</td>
</tr>
<tr>
<td>ACToR “Chemical/Industrial Process use with no Consumer use”</td>
<td>Chemical substances and products in industrial manufacturing processes that are not used in consumer products. Does not include food or pharmaceuticals</td>
<td>14   282</td>
</tr>
<tr>
<td>ACToR UseDB “Pesticide Inert use”</td>
<td>Secondary (i.e., non-active) ingredients in a pesticide which serve a purpose other than repelling pests. Pesticide use of these ingredients is known due to more stringent reporting standards for pesticide ingredients, but many of these chemicals appear to be also used in consumer products</td>
<td>16   816</td>
</tr>
<tr>
<td>ACToR “Pesticide Active use”</td>
<td>Active ingredients in products designed to prevent, destroy, repel, or reduce pests (e.g., insect repellants, weed killers, and disinfectants).</td>
<td>76   877</td>
</tr>
<tr>
<td>TSCA IUR 2006 Total Production Volume</td>
<td>Sum total (kg/year) of production of the chemical from all sites that produced the chemical in quantities of 25,000 pounds or more per year. If information for a chemical is not available, it is assumed to be produced at &lt;25,000 pounds per year.</td>
<td>106  7784</td>
</tr>
</tbody>
</table>

Wambaugh et al., (2014)
Predictors Do Not Vary Between Groups

- The vertical lines indicate the 95% credible interval across the 1500 different exposure scenarios inferred from the NHANES urine data.

- SHEDS-HT (Isaacs et al., 2014) should help explain some remaining NHANES variability.

Wambaugh et al., (2014)
Calibrated Exposure Predictions for 7968 Chemicals

Wambaugh et al., (2014)
• We focus on the median and upper 95% predictions because the lower 95% is below the NHANES limits of detection (LoD)

• Dotted lines indicate 25%, median, and 75% of the LoD distribution
• Chemicals currently monitored by NHANES are distributed throughout the predictions
• Chemicals with the first and ninth highest 95% limit are monitored by NHANES

Wambaugh et al., (2014)
Calibrated Exposure Predictions for 7968 Chemicals

- The grey stripes indicate the 4182 chemicals with no use indicated by ACToR UseDB for any of the four use category heuristics

Wambaugh et al., (2014)
High Throughput Risk Prioritization

ToxCast Bioactivity Converted to mg/kg/day with HTTK

in vitro Activities

Exposure Prediction (Median and Upper 95%)

ExpoCast Exposure Predictions

Prioritization as in Wetmore et al. (2012) Bioactivity, Dosimetry, and Exposure Paper
A Closer Look at Bisphenol A

- LaKind and Naiman (2011) Estimated Exposure to BPA from NHANES data in ng/kgBW/day):

<table>
<thead>
<tr>
<th>Demographic</th>
<th>LaKind and Naiman (2011)</th>
<th>ExpoCast Geometric Mean Median</th>
<th>ExpoCast Geometric Mean Upper 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>35.1</td>
<td>25.0</td>
<td>2193</td>
</tr>
<tr>
<td>Age 6-11y</td>
<td>54</td>
<td>63</td>
<td>4984</td>
</tr>
<tr>
<td>Age 12-19y</td>
<td>48</td>
<td>59</td>
<td>5169</td>
</tr>
<tr>
<td>Age 20-39y*</td>
<td>38.5</td>
<td>57</td>
<td>6056</td>
</tr>
<tr>
<td>Age 40-59y*</td>
<td>28.9</td>
<td>57</td>
<td>6056</td>
</tr>
<tr>
<td>Age &gt;=60y</td>
<td>27.3</td>
<td>66</td>
<td>84221</td>
</tr>
<tr>
<td>Male</td>
<td>39.6</td>
<td>38</td>
<td>3132</td>
</tr>
<tr>
<td>Female</td>
<td>31.2</td>
<td>12</td>
<td>1125</td>
</tr>
</tbody>
</table>

- CPCPdb (Goldsmith et al., 2014): 1797 unique chemicals mapped to 8921 consumer products, but no Bisphenol A

*ExpoCast makes single prediction for Age 20-59y
A Closer Look at Triclosan

- EPA Triclosan Occupational and Residential Exposure Assessment (2008) µg/kg BW/d exposures:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>2.5</td>
<td>2.9</td>
<td>2.9</td>
<td>4.5</td>
<td>0.0012</td>
<td>0.085</td>
</tr>
<tr>
<td>Age 6-11</td>
<td>1.6</td>
<td>1.9</td>
<td>1.7</td>
<td>2.4</td>
<td>0.0079</td>
<td>0.17</td>
</tr>
<tr>
<td>Age 12-19</td>
<td>2.7</td>
<td>3.2</td>
<td>4.1</td>
<td>6.2</td>
<td>0.0015</td>
<td>0.11</td>
</tr>
<tr>
<td>Age 20-59</td>
<td>2.9</td>
<td>3.2</td>
<td>3.0</td>
<td>4.7</td>
<td>0.0015</td>
<td>0.11</td>
</tr>
<tr>
<td>Age &gt;=60</td>
<td>1.9</td>
<td>2.2</td>
<td>2.1</td>
<td>3.3</td>
<td>0.002</td>
<td>0.083</td>
</tr>
<tr>
<td>Male</td>
<td>3.1</td>
<td>3.8</td>
<td>3.6</td>
<td>5.6</td>
<td>0.0011</td>
<td>0.074</td>
</tr>
<tr>
<td>Female</td>
<td>2.0</td>
<td>2.1</td>
<td>2.1</td>
<td>3.4</td>
<td>0.0016</td>
<td>0.11</td>
</tr>
</tbody>
</table>

- Triclosan exposures underestimated by ExpoCast because most pesticide active exposures are significantly lower than exposures for other chemical classes – SHEDS-HT should help
SEEM Evolution

Model and Predictors

1st Gen
- USEtox
- RAIDAR
- Near Field / Far Field
- Production Volume

2nd Gen
- Use Categories
- Production Volume
- Phys-Chem Properties

3rd Gen
- SHEDS-HT
- Literature Models
- CPcat Database
- Pesticide REDs

Calibration/Evaluation Data
- NHANES Urine Data
- NHANES Urine and Blood Data

SEEM Conclusion
• Existing complex fate and transport models have low correlation to measured exposures
  • Near field factor most important
• Simple, readily available data
  • Better correlation to measured exposures
  • Similar predictions across demographics
• Isaacs et al. (2014) developed SHEDS-HT and predicted consumer product-driven exposures for 2507 chemicals
  • SEEM analysis in progress
Better Models and Data Should Reduce Uncertainty

Uncertainty/Variability of NHANES Biomonitoring

~10% Far field (Industrial) Releases

~60% Indoor / Consumer Use

Consumer product database and two new near field models in 2014
Data Inhomogeneity

7968 Chemicals

- Physico-chemical
- ACToR UseDB
- CPCPdb
- SHEDS-HT Dietary
- SHEDS-HT Residues
- Pesticide Documents
Conclusions

• We identify those HTE factors that correlate with the NHANES data and estimate uncertainty

• The calibrated meta-model can estimate relative levels of chemical exposures for 7968 chemicals
  – This includes thousands of chemicals with no other data on human exposure
  – Same factors are predictive ($R^2 \sim 0.5$) across demographics characterized by NHANES

• Different demographics have different mean (overall) exposures:
  – There are demographic-specific aspects not currently described by available HTE factors

• Upcoming analysis:
  – Augment heuristics with calibrations of new mechanistic HT models for exposure from consumer use and indoor environment (e.g., SHEDS-HT)
  – Develop new data sources with additional chemical descriptors (e.g., CPcatDB)
  – Should help decrease uncertainties and increase confidence in extrapolation
EPA Office of Research and Development

NCCT
Chris Grulke
Richard Judson
Thomas Knudsen
Chantel Nicolas*
Robert Pearce*
James Rabinowitz
Woody Setzer
Imran Shah
Rusty Thomas

NRMRL
Jane Bare
Xiaoyu Liu

NERL
Craig Barber
Dalizza Colon
Kathie Dionisio*
Peter Egeghy
Kim Gaetz
Kristin Isaacs
Haluk Ozkaynak
Julia Rager*
Mark Strynar
Jon Sobus
Mike Tornero-Velez
Dan Vallero

NHEERL
Jane Ellen Simmons
Marina Evans
Mike Hughes

*Trainees

Arnot Research and Consulting
Jon Arnot
Chemical Computing Group
Rocky Goldsmith
Environmental Protection Agency
Alicia Frame
Hamner Institutes
Barbara Wetmore
Cory Strope
Indiana University
James Sluka
Michigan State University
Jade Mitchell
National Institute for Environmental Health Sciences (NIEHS)
Nisha Sipes
Kyla Taylor
Kristina Thayer
Netherlands Organisation for Applied Scientific Research (TNO)
Sieto Bosgra
North Carolina State University
Anran Wang
University of California, Davis
Deborah Bennett
University of Michigan
Olivier Jolliet
University of North Carolina, Chapel Hill
Alexander Sedykh *
Alex Tropsha


