

### **EPA's Rapid Exposure and Dosimetry Project**

John Wambaugh National Center for Computational Toxicology Office of Research and Development U.S. Environmental Protection Agency wambaugh.john@epa.gov

**September 15, 2017** 

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



### Chemical Safety for Sustainability (CSS) Rapid Exposure and Dosimetry (RED) Project Co-Leads Kristin Isaacs and John Wambaugh

#### **NCCT**

Chris Grulke
Greg Honda\*

Richard Judson

Andrew McEachran\*

Robert Pearce\*

**Ann Richard** 

Parichehr

Saranjampour\*

Risa Sayre\*

**Woody Setzer** 

Rusty Thomas

John Wambaugh

**Antony Williams** 

#### NRMRL

Yirui Liang\* Xiaoyu Liu

#### **NHEERL**

Linda Adams

Christopher

**Ecklund** 

Marina Evans

Mike Hughes

Jane Ellen

Simmons

\*Trainees

#### **NERL**

Craig Barber

Namdi Brandon\*

Peter Egeghy

Hongtai Huang\*

Brandall Ingle\*

**Kristin Isaacs** 

**Seth Newton** 

**Katherine Phillips** 

**Paul Price** 

Jeanette Reves\*

Jon Sobus

John Streicher\*

Mark Strynar

Mike Tornero-Velez

Elin Ulrich

Dan Vallero

Barbara Wetmore

#### **Collaborators**

**Arnot Research and Consulting** 

Jon Arnot

**Battelle Memorial Institute** 

**Anne Louise Sumner** 

**Anne Gregg** 

**Chemical Computing Group** 

**Rocky Goldsmith** 

**National Institute for Environmental Health** 

**Sciences (NIEHS) National Toxicology Program** 

Mike Devito

Steve Ferguson

Nisha Sipes

**Netherlands Organisation for Applied Scientifi** 

Research (TNO)
Sieto Bosgra

**Research Triangle Institute** 

Timothy Fennell

ScitoVation

Harvey Clewell
Chantel Nicolas

**Silent Spring Institute** 

**Robin Dodson** 

**Southwest Research Institute** 

Alice Yau

Kristin Favela

**Summit Toxicology** 

Lesa Aylward

**Tox Strategies** 

Caroline Ring

**University of California, Davis** 

**Deborah Bennett** 

**University of Michigan** 

Olivier Jolliet

**University of North Carolina, Chapel Hill** 

Alex Tropsha

University of Texas, Arlington

Hyeong-Moo Shin

#### **Lead CSS Matrix Interfaces:**

John Kenneke (NERL)
John Cowden (NCCT)



### 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

 Park et al. (2012): At least 3221 chemicals in humans, many appear to be exogenous

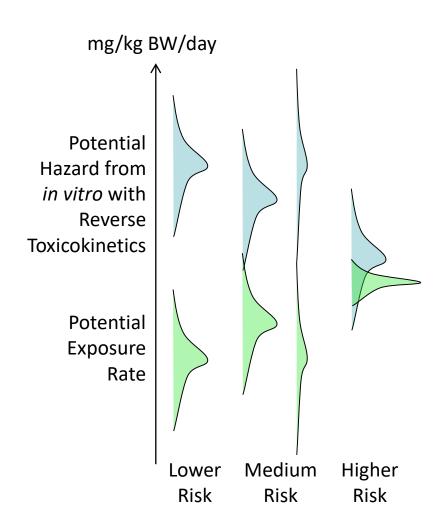


November 29, 2014



## High Throughput Risk Prioritization

- High throughput risk prioritization needs:
  - high throughput hazard characterization (from HTT project)
  - high throughput exposure forecasts
  - 3. high throughput **toxicokinetics** (*i.e.*, dosimetry)





# High-Throughput Bioactivity

■ **Tox21**: Examining >10,000 chemicals using ~50 assays intended to identify interactions with biological pathways (Schmidt, 2009)



EPA Toxicity Forecaster (ToxCast):

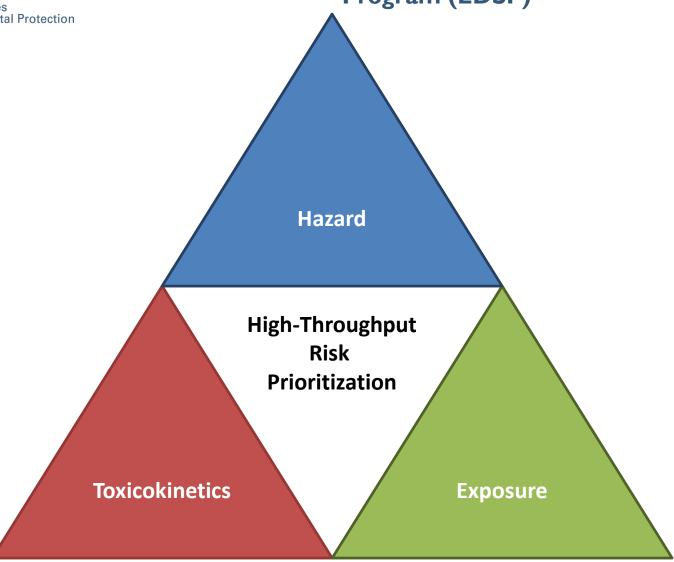
For a subset (>3000) of Tox21 chemicals run >1000 additional assay endpoints (Judson et al., 2010)

Most assays conducted in dose-response format (identify 50% activity concentration – AC50 – and efficacy if data described by a Hill function)

http://actor.epa.gov/dashboard/



Application to U.S. EPA Endocrine Disruptor Screening Program (EDSP)

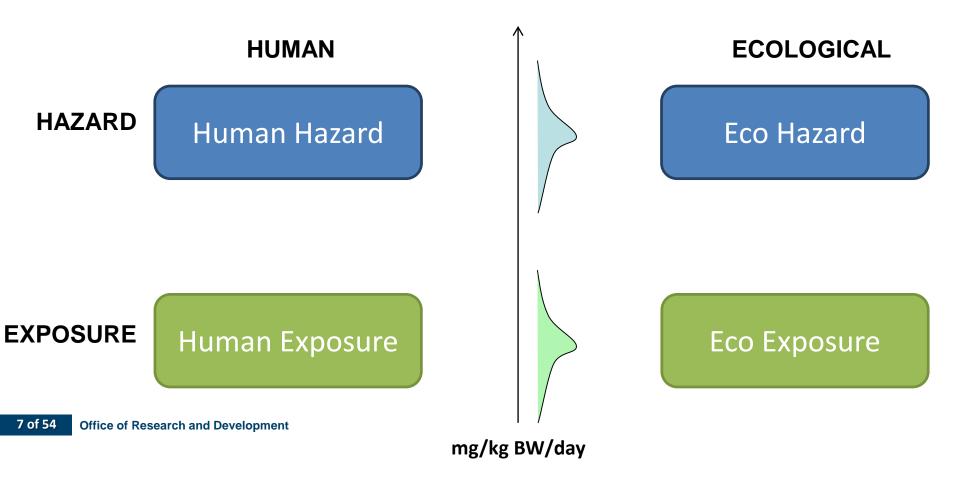




### High Throughput Chemical Risk Prioritization

Prioritization as in Wetmore *et al.* (2015)

July and December 2014 FIFRA Scientific Advisory Panels reviewed research as it applies to the Endocrine Disruptor Screening Program

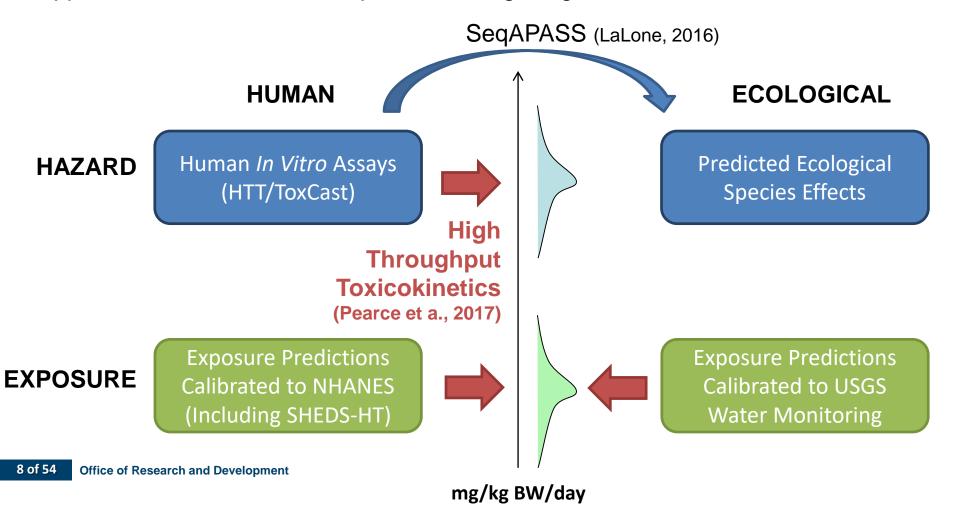




### High Throughput Chemical Risk Prioritization

Prioritization as in Wetmore *et al.* (2015)

July and December 2014 FIFRA Scientific Advisory Panels reviewed research as it applies to the Endocrine Disruptor Screening Program





### High Throughput Risk Prioritization in Practice

mg/kg bw/day

ToxCast-derived
Receptor Bioactivity
Converted to
mg/kg/day with
HTTK

ExpoCast Exposure Predictions

Near Field Far Field

**ToxCast Chemicals** 

December, 2014 Panel:

"Scientific Issues Associated with Integrated Endocrine Bioactivity and Exposure-Based Prioritization and Screening"

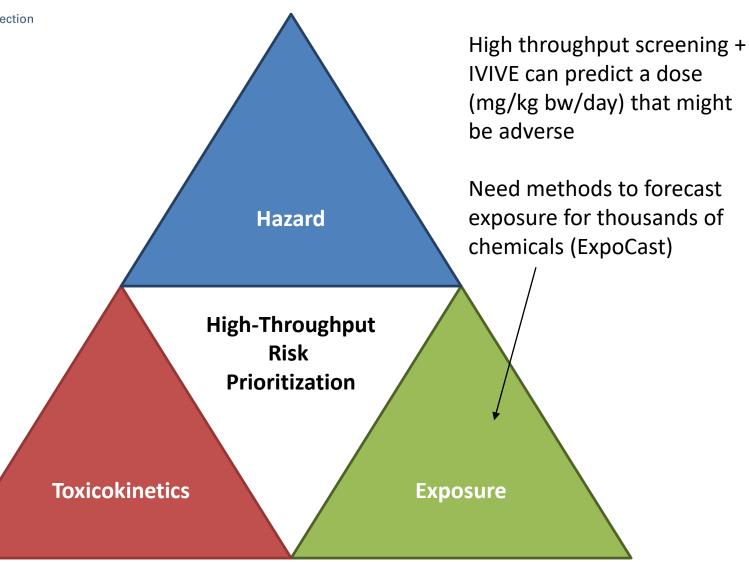
ToxCast: Toxicity Forecaster

ExpoCast: Exposure Forecaster

Rapid Exposure and Dosimetry Project provides ExpoCast research

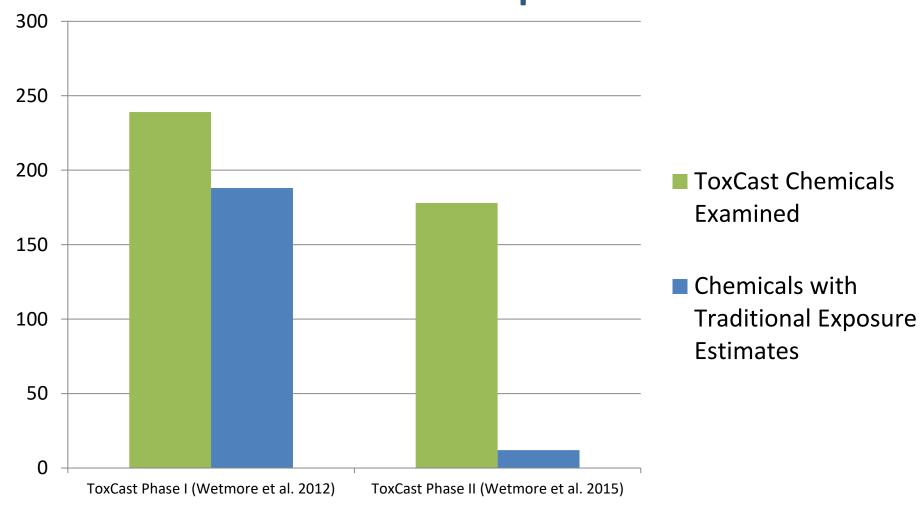


### **High Throughput Exposure**





# The Need for High Throughput Exposure



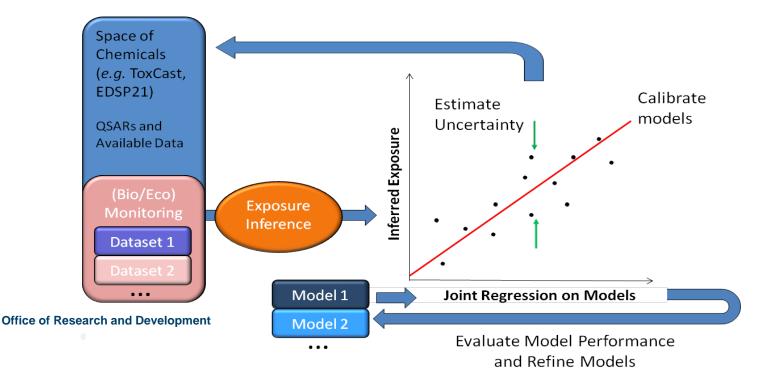
Egeghy et al. (2012) – Most chemicals lack exposure data



12 of 54

### Consensus Exposure Predictions with the SEEM Framework

- We incorporate multiple models into consensus predictions for 1000s of chemicals within the
   Systematic Empirical Evaluation of Models (SEEM) framework (Wambaugh et al., 2013, 2014)
- We evaluate/calibrate predictions with available monitoring data
- This provides information similar to a sensitivity analysis: What models are working?
   What data are most needed? This is an iterative process.
- To date we have relied on median U.S. population exposure rates only





# **Exposures Inferred from NHANES**

 Annual survey, data released on 2-year cycle.

National Health and Nutrition Examination Survey

- Different predictive models provide different chemicalspecific predictions
  - Some models may do a better job form some chemical classes than others overall, so we want to evaluate performance against monitoring data
- Separate evaluations can be done for various demographics



### **Heuristics of Exposure**

Wambaugh et al. (2014)

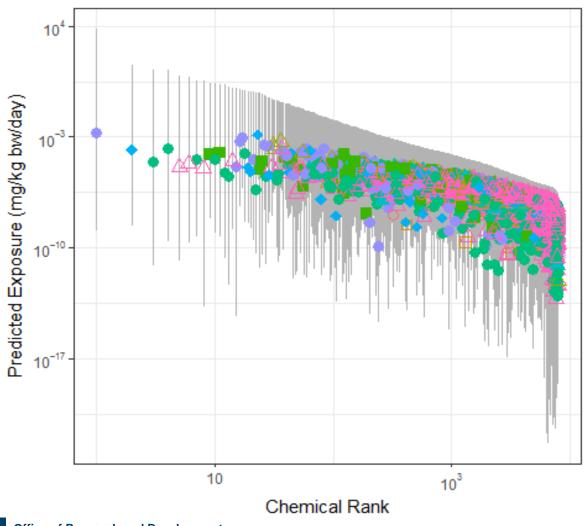
Five descriptors explain roughly 50% of the chemical to chemical variability in median NHANES exposure rates

Same five predictors work for all NHANES demographic groups analyzed – stratified by age, sex, and body-mass index:

- Industrial and Consumer use
- Pesticide Inert
- Pesticide Active
- Industrial but no Consumer use
- Production Volume



# Human Exposure Predictions for 134,521 Chemicals



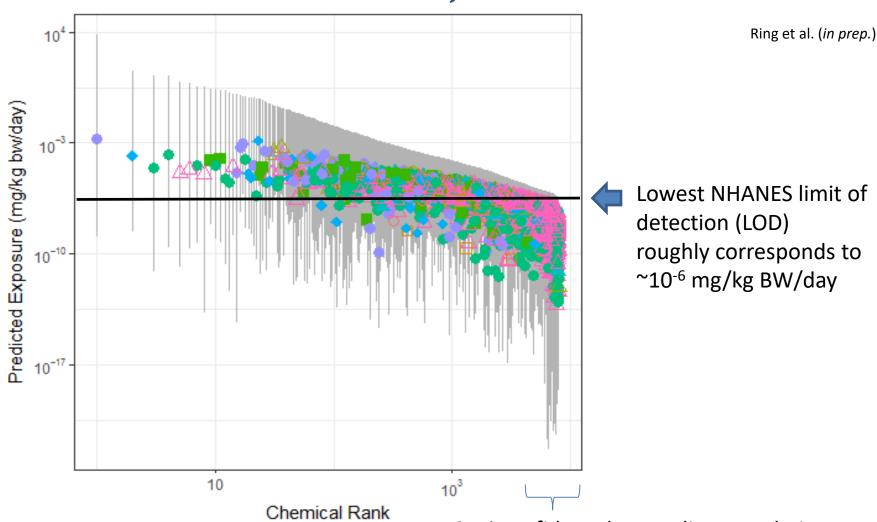
Ring et al. (in prep.)

#### Pathway

- Dietary
- Dietary, Industrial
- Dietary, Residential
- Dietary, Residential, Industrial
- Industrial
- Pesticide
- Residential
- Residential, Industrial
- Residential, Pesticide
- △ Unknown



# Human Exposure Predictions for 134,521 Chemicals



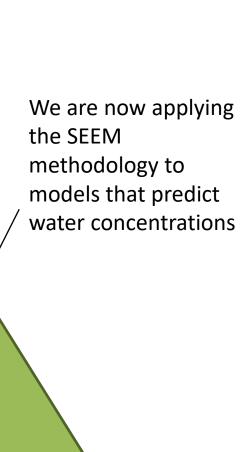
# **Environmental Protection** Agency **High-Throughput**

#### **Water Concentrations**

Hazard

Risk

**Prioritization** 



**Toxicokinetics** 

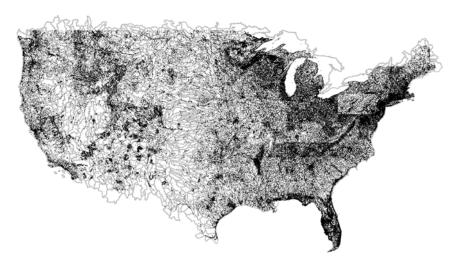
**Exposure** 



# USGS National Water Quality Assessment (NAWQA) Data

Watersheds

https://www.waterqualitydata.us/portal/



1984 – 2014

Aggregated by season

**HUC Levels** 

**National** 

#### **Surface Water Sampling Sites**

- > 600,000 surface water sites in lower 48
- > 700 individual chemicals

GPS, date, and time stamps

LOD indication

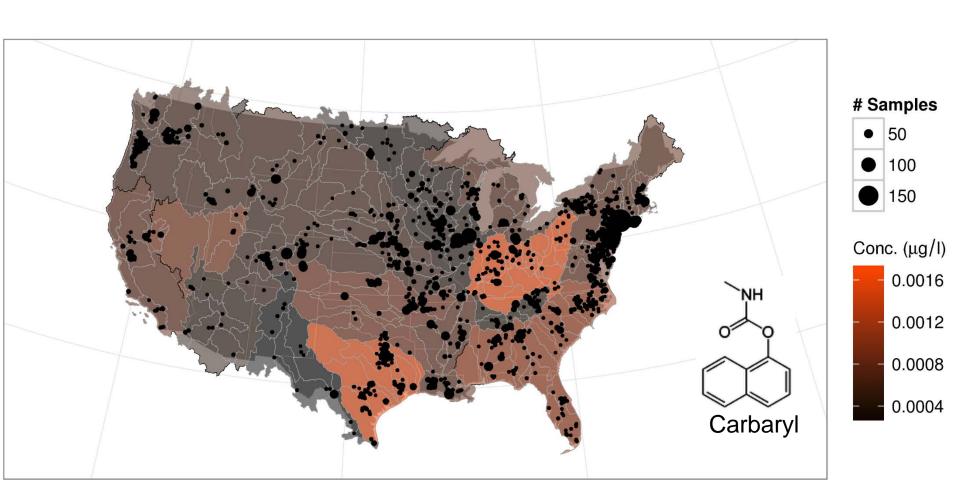
Region (HUC2) n = 18

Sub-region (HUC4) n = 196

HUC = hydrological unit



### Regional Watersheds (HUC2)



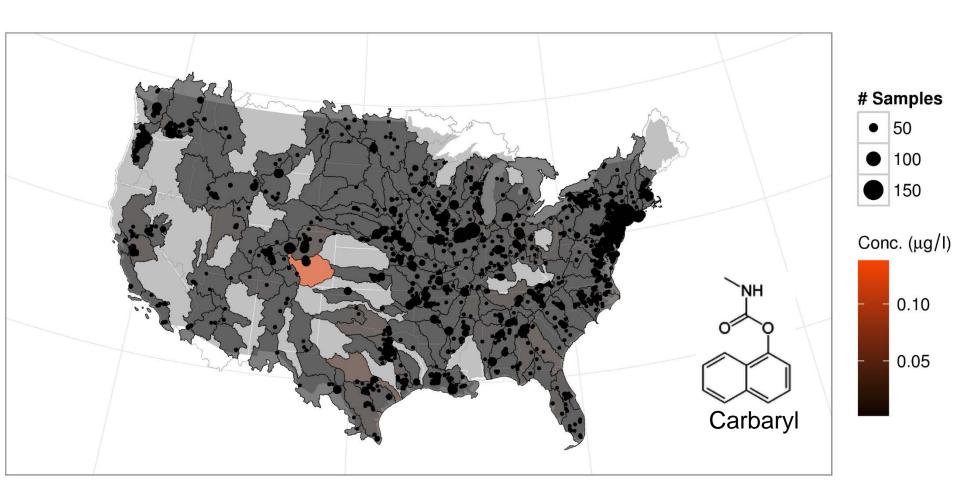
19 of 54

Office of Research and Development

Setzer et al., (in prep)



# Sub-Regional Watersheds (HUC4)



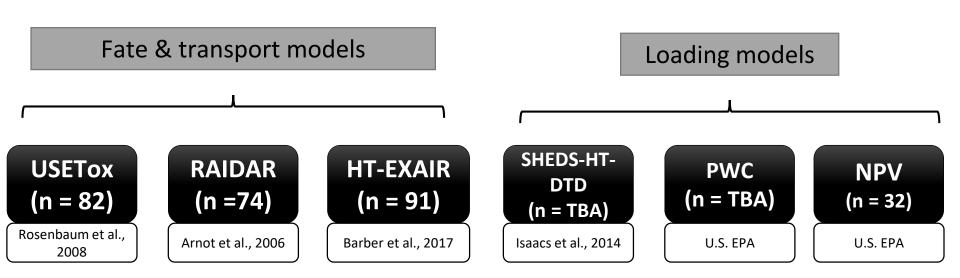
20 of 54

Office of Research and Development

Setzer et al., (in prep)



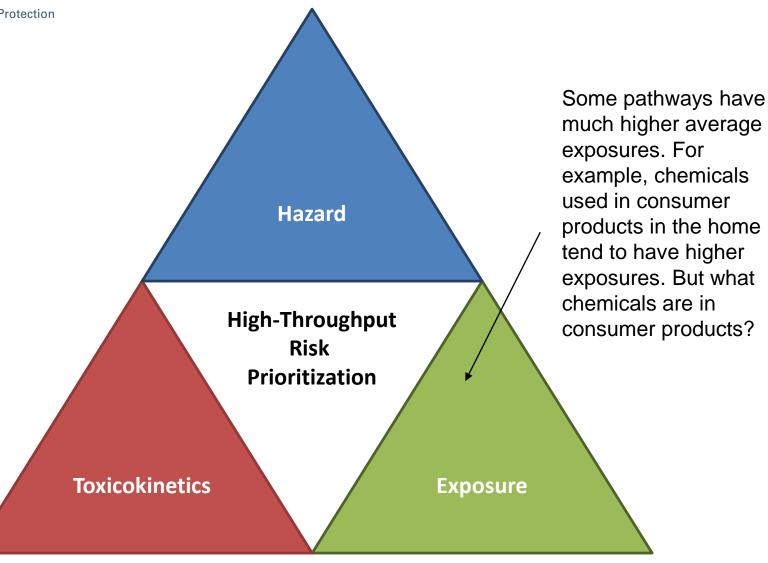
# Predicting Water Concentrations for Thousands of Chemicals



 $\mu$  = geometric mean water concentrations log ( $\mu$ ) = fate and transport models \* loading models

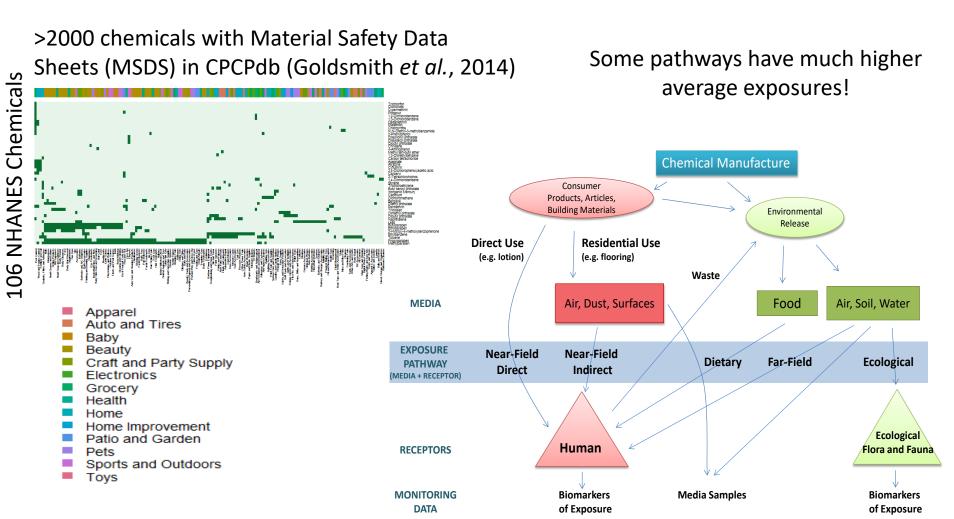
# United States Environmental Protection Agency

### **Idenifying Exposure Pathways**





### Chemical Use Identifies Relevant Pathways



23 of 54

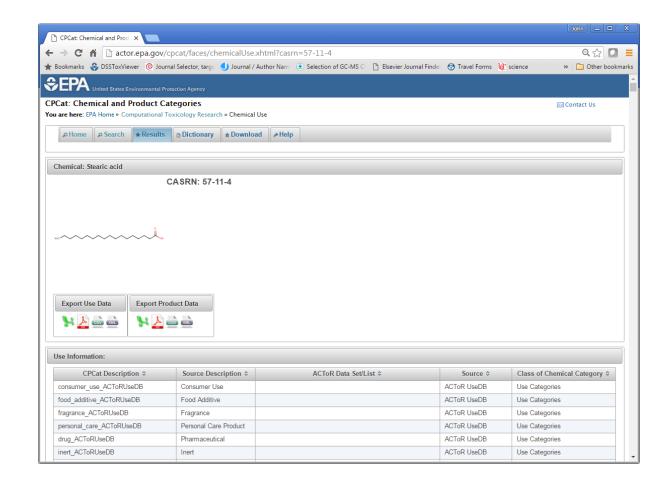
Office of Research and Development

Near field sources have been known to be important at least since 1987 – see Wallace, et al.



### CPdat: Chemical Use Information for ~30,000 Chemicals

- Chemical-Product database (CPdat) maps many different types of use information and ontologies onto each other
- Includes CPCPdb
   (Goldsmith, et al., 2014)
   with information on
   ~2000 products from
   major retailors
- Largest single database has coarsest information: ACToR UseDB





### **Predicting Exposure**

- EPA's public CPdat (<a href="http://actor.epa.gov/cpcat/">http://actor.epa.gov/cpcat/</a>) includes every chemical safety sheet from a major U.S. retailer (>2000 chemicals) but there are many thousands of other chemicals (Goldsmith et al, 2015)
- We use applied statistics, including machine learning techniques, to learn from the data we have to fill in the gaps (Wambaugh et al., 2014, Isaacs et al., 2016, Phillips et al., 2017)
  - This is similar to how Netflix can guess how much you will like a movie



A \$1 Million Research Bargain for Netflix, and Maybe a Model for Others

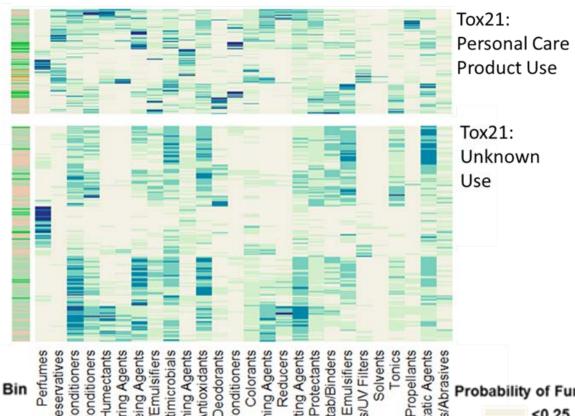
By STEVE LOHR Published: September 21, 2009

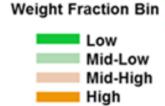




### **Predicting Chemical Constituents**

- CPCPdb does not cover every chemical-product combination (~2000 chemicals, but already >8000 in Tox21)
- We can predict functional use and weight fraction for thousands of chemicals



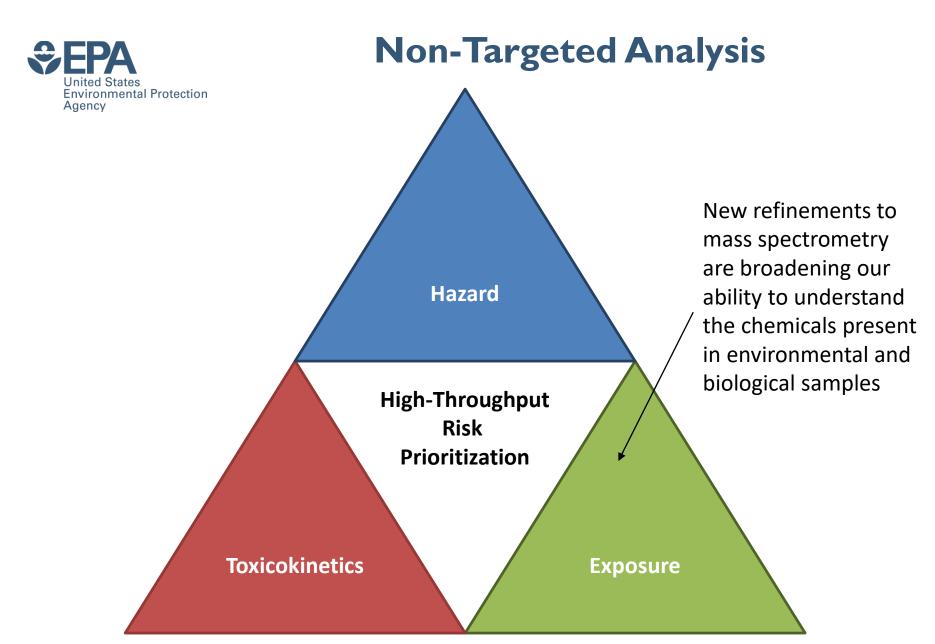


Skin Conditioners Emollients/Skin

Probability of Function < 0.25

0.25-0.5 0.5-0.75 0.75-0.9 >0.9

Isaacs *et al.* (2016)





### Non-Targeted and Suspect-Screening Analysis

- Models present one way forward, but new analytic techniques may also allow insight in to chemicals composition of products and the greater environment
- EPA is coordinating a comparison of nontargeted screening workflows used by leading academic and government groups (led by Jon Sobus and Elin Ulrich)
  - Examining house dust, human plasma, and silicone wristbands (O'Connell, et al., 2014)
  - Similar to NORMAN Network (Schymanski et al., 2015) analysis of water
- Published analysis on house dust (Rager et al., 2016)

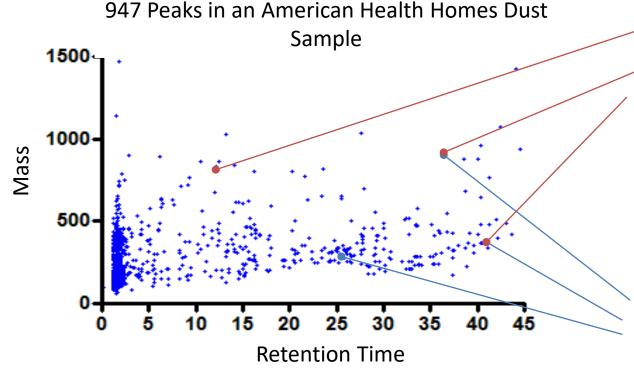


"I'm searching for my keys."

100 consumer products from a major U.S. retailer were analyzed, tentatively identifying 1,632 chemicals, 1,445 which were not in EPA's database of consumer product chemicals (Phillips *et al.*, *submitted*)



### **Suspect Screening in House Dust**



Each peak corresponds to a chemical with an accurate mass and predicted formula:

$$C_{17}H_{19}NO_3$$

Multiple chemicals can have the same mass and formula:

Is chemical A present, chemical B, both, or some other chemical (neither)?

We are expanding our reference libraries using ToxCast chemicals to enable greater numbers and better accuracy of confirmed chemicals



### Appropriate Skepticism for Non-Targeted Analysis and Suspect Screening

"As chemists we are obliged to accept the assignment of barium to the observed activity, but as nuclear chemists working very closely to the field of physics we cannot yet bring ourselves to take such a drastic step, which goes against all previous experience in nuclear physics. It could be, however, that a series of strange coincidences has misled us."

Hahn and Strassmann (1938)



### Appropriate Skepticism for Non-Targeted Analysis and Suspect Screening

"As chemists we are obliged to accept the assignment of barium to the observed activity, but as nuclear chemists working very closely to the field of physics we cannot yet bring ourselves to take such a drastic step, which goes against all previous experience in nuclear physics. It could be, however, that a series of strange coincidences has misled us."

Hahn and Strassmann (1938)

1944 Nobel Prize in Chemistry for "discovery of the fission of heavy nuclei"



 $Log_{10}(\mu g/g)$ 

The chemicals found in a cotton shirt





 $Log_{10}(\mu g/g)$ 

Chemicals that are present



Chemicals that are absent (but found in other products)



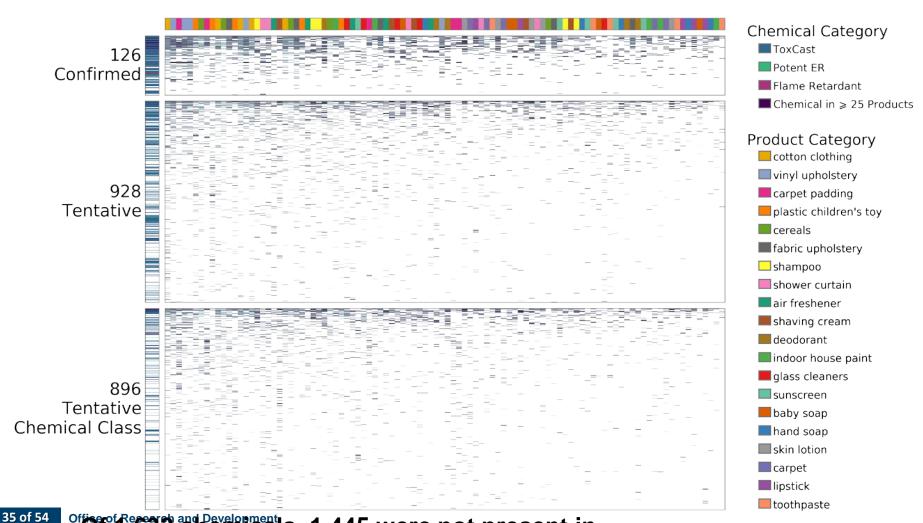
 $Log_{10}(\mu g/g)$ 

The chemicals found in a cotton shirt





 $Log_{10}(\mu g/g)$ 



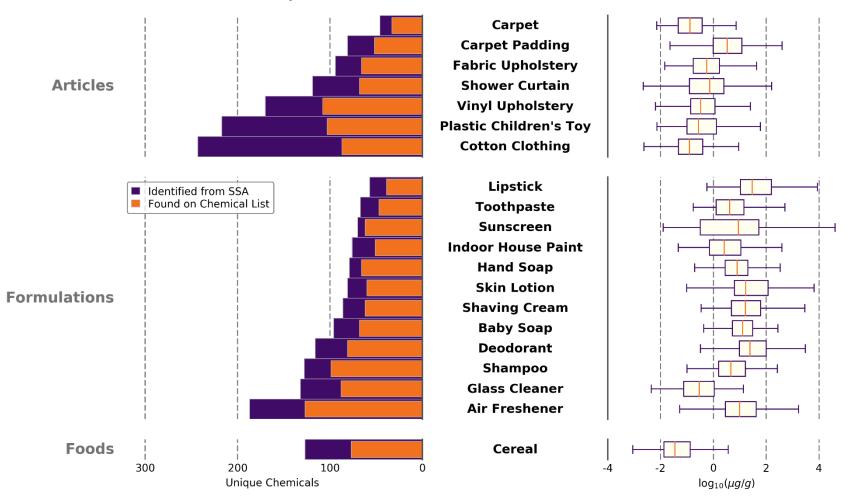
of 1,632 chemicals, 1,445 were not present in our database from the major retailer (CPCPdb)

Phillips et al. (submitted)



### **Product Scan Summary**

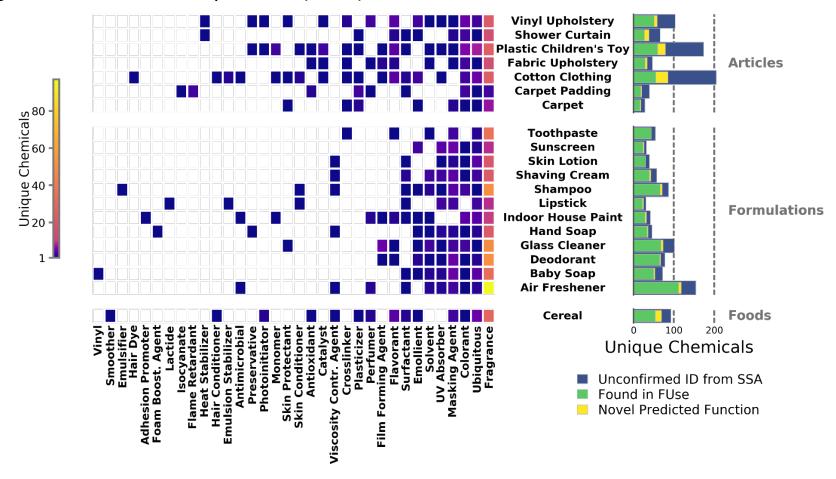
Of 1,632 chemicals confirmed or tentatively identified, 1,445 were not present in CPCPdb





### **Predicting Chemical Function**

#### Using the methods of Phillips *et al.*, (2017):

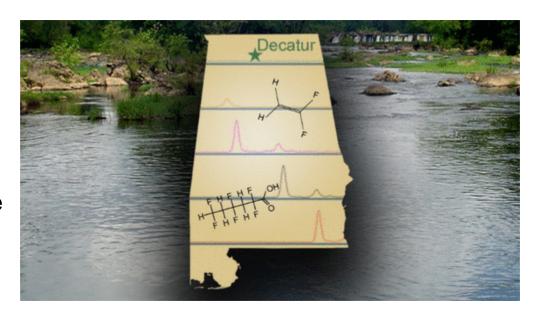


**Chemical Function** 



### **Analysis of Drinking Water**

High resolution mass spectrometry was used to investigate the occurrence and identity of replacement fluorinated compounds in surface water and sediment of the Tennessee River near Decatur, Alabama



A series of nine polyfluorinated carboxylic acids was discovered



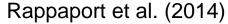
# Caveats to Non-Targeted Screening

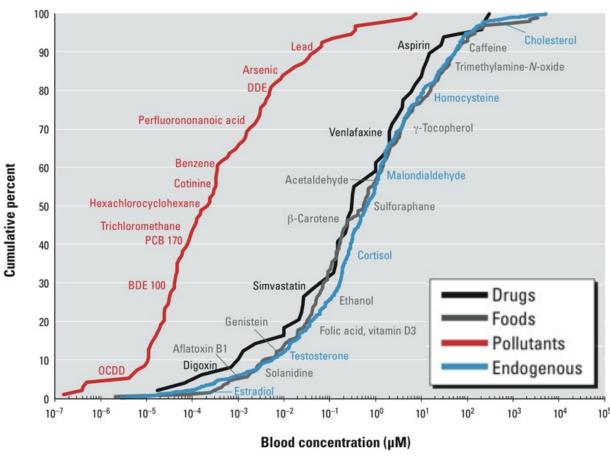
- Chemical presence in an object does not mean that exposure occurs
- Only some chemical identities are confirmed, most are tentative
  - Can use formulation predictor models as additional evidence
- Chemical presence in an object does not necessarily mean that it is bioavailable
  - Can build emission models
- Small range for quantitation leads to underestimation of concentration
- Product de-formulation caveats:
  - Samples are being homogenized (e.g., grinding) and are extracted with a solvent (dichloro methane, DCM)
  - Only using one solvent (DCM, polar) and one method GCxGC-TOF-MS
  - Varying exposure intimacy, from carpet padding to shampoo to cereal
- Exposure alone is not risk, need hazard data



### **Expanded Biomonitoring**

- Moving beyond NHANES chemicals
  - Non-targeted analysis of blood may be possible
  - Not just a matter of sensitivity, must also "filter out" endogenous, food, and drug chemicals





# United States Environmental Protection Agency

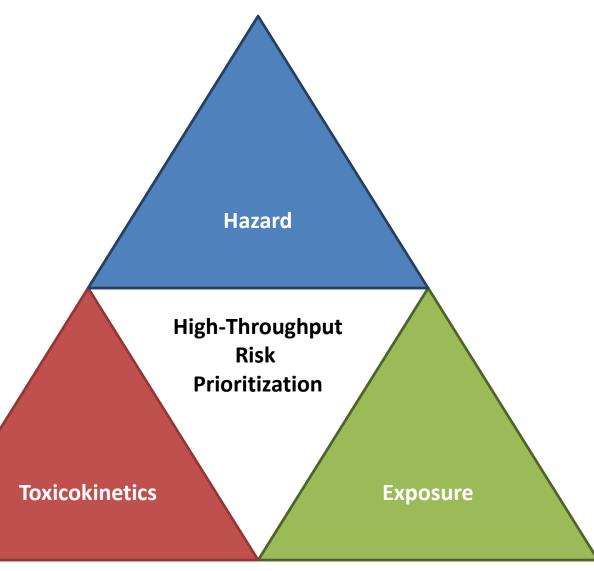
We want to perform

extrapolation (IVIVE)

of ToxCast activities

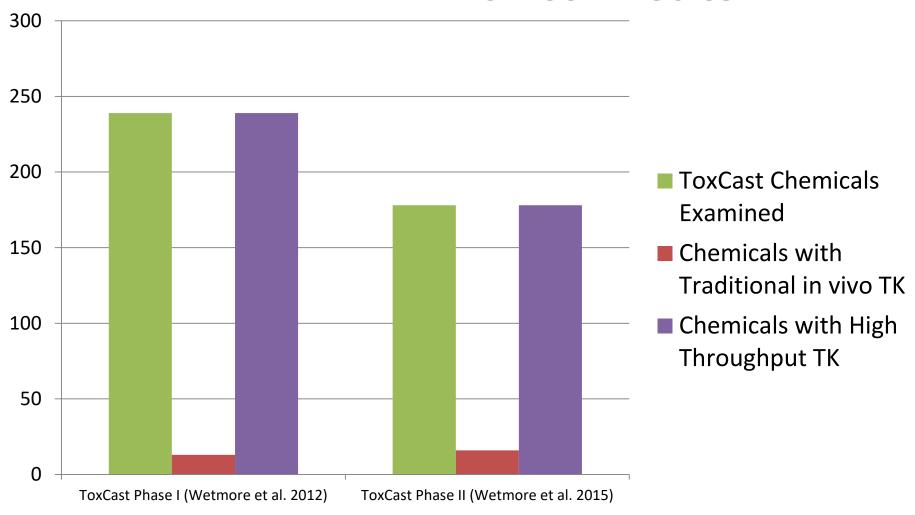
in vitro-in vivo

#### **Toxicokinetics for IVIVE**





# The Need for In Vitro Toxicokinetics



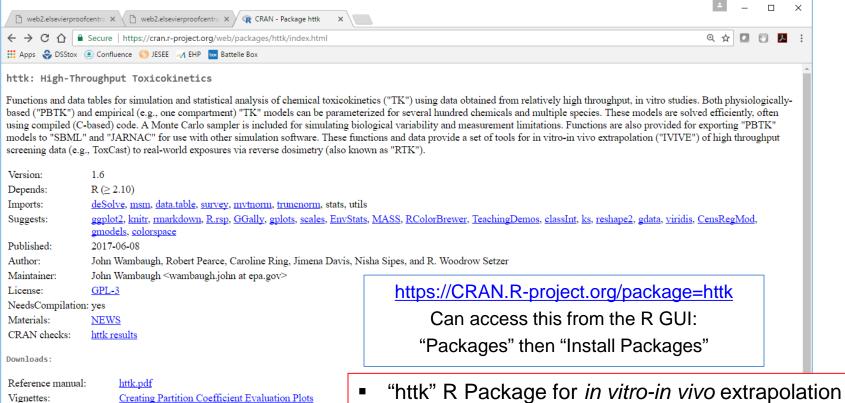
42 of 54

Office of Research and Development

 Studies like Wetmore et al. (2012, 2015), addressed the need for TK data using in vitro methods

### **High Throughput Toxicokinetics**





Age distributions

Global sensitivity analysis

Global sensitivity analysis plotting Height and weight spline fits and residuals Hematocrit spline fits and residuals

Plotting Css95

Serum creatinine spline fits and residuals

Generating subpopulations

Evaluating HTTK models for subpopulations

Generating Figure 2 Generating Figure 3

Plotting Howgate/Johnson data

AER plotting

Virtual study populations

httk: R Package for High-Throughput Toxicokinetic

- and PBTK
- 553 chemicals to date
- 100's of additional chemicals being studied
- Pearce et al. (2017) provides documentation and examples
- Built-in vignettes provide further examples of how to use many functions

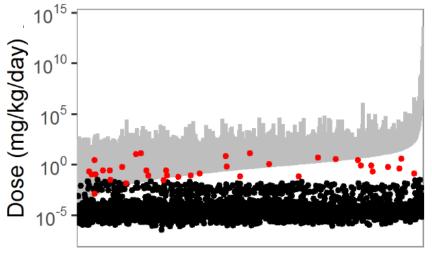
43 of 54



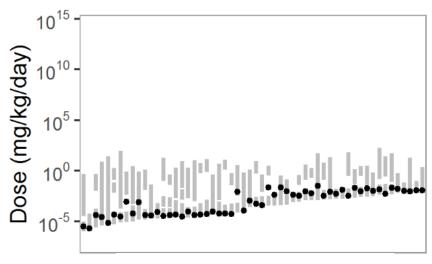
## Using HTTK Predicted C<sub>max</sub> for Risk Prioritization



Screening for toxicity has blind spots and exposure forecasts are highly uncertain, yet:



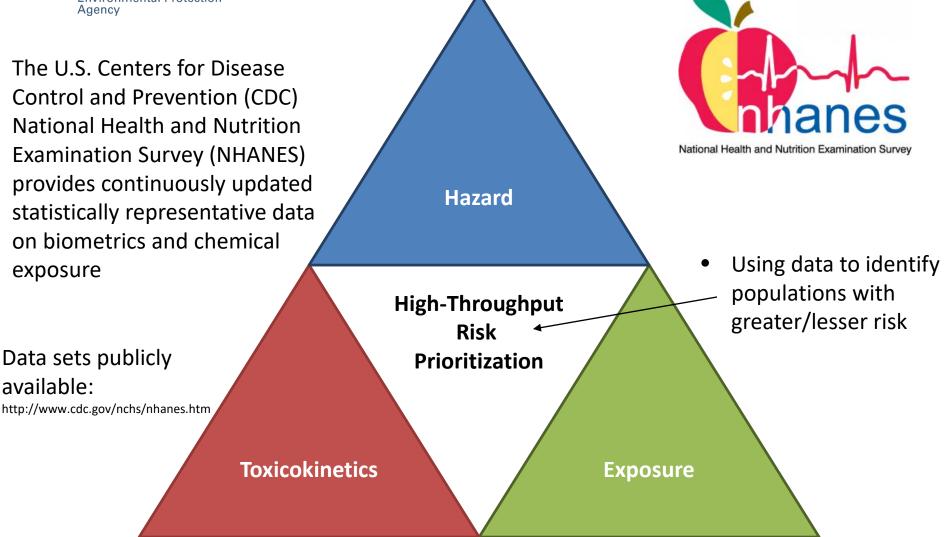
Doses ranges for all 3925 Tox21 compounds eliciting a 'possible'-to-'likely' human *in vivo* interaction alongside estimated daily exposure



56 compounds with potential *in vivo* biological interaction at or above estimated environmental exposures



### Further Analyzing the CDC NHANES Data



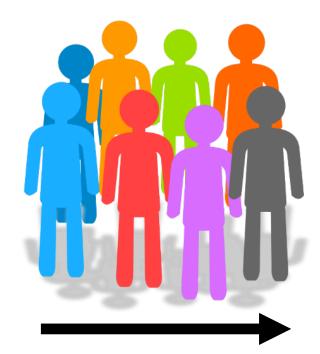


### Population simulator for HTTK

Correlated Monte Carlo sampling of physiological model parameters

## **Sample NHANES** quantities

Sex
Race/ethnicity
Age
Height
Weight
Serum creatinine



## Regression equations from literature (+ residual marginal variability)

### **Predict** physiological quantities

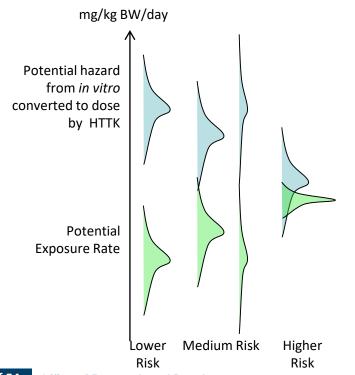
Tissue masses
Tissue blood flows
GFR (kidney function)
Hepatocellularity

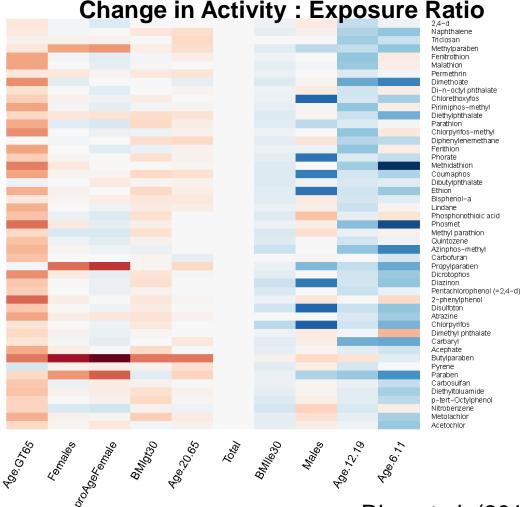


### Toxicokinetic IVIVE: Convert HTS μM to mg/kg/day

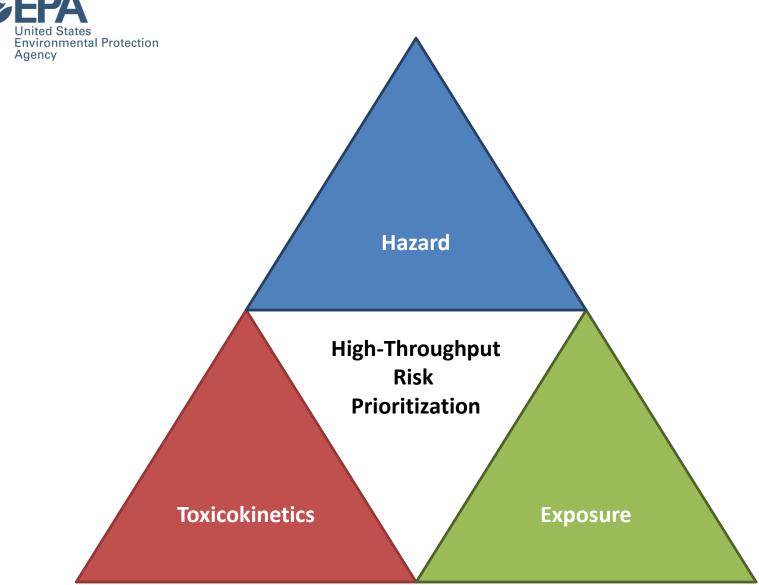
 We use HTTK to calculate margin between bioactivity and exposure for specific populations

Change in Risk





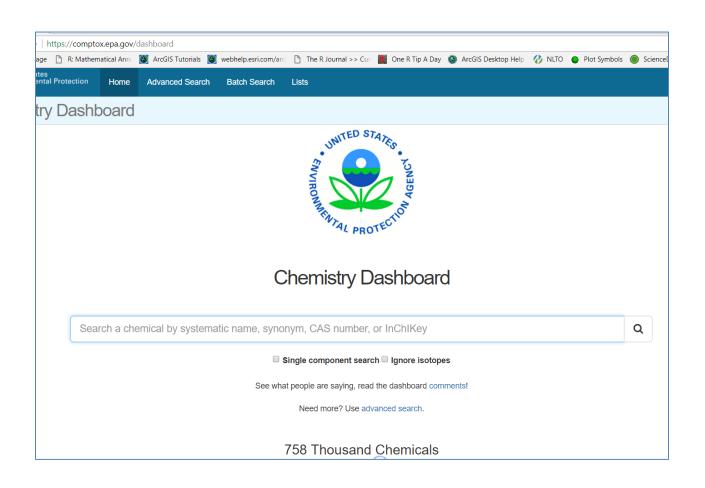
#### **Public Chemical Assessment Tools from EPA ORD**





### A Google for Chemicals

#### http://comptox.epa.gov/dashboard/



Examples:

"Stearic Acid"

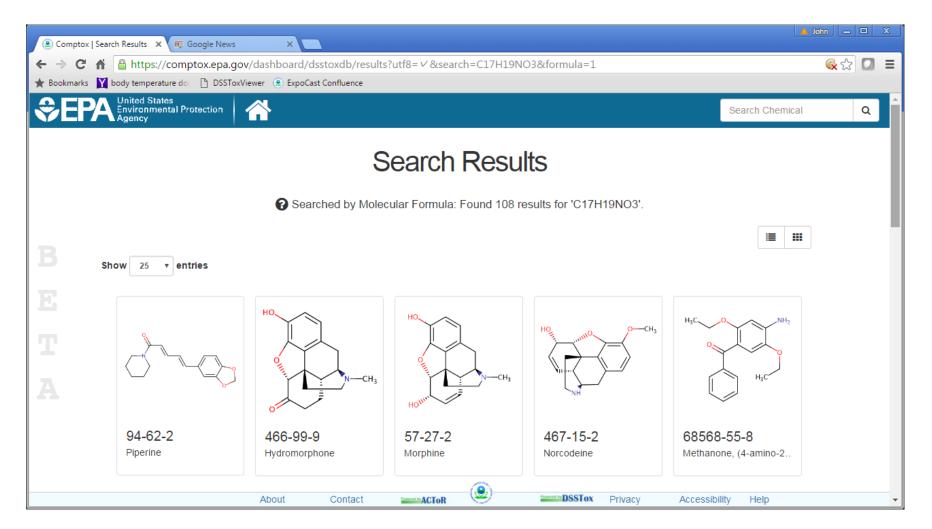
"Bisphenol A"

"C17H19NO3"



### A Google for Chemicals

http://comptox.epa.gov/dashboard/





## Public Chemical Assessment Tools from EPA ORD

Dashboards: Chemistry Dashboard (one stop shop):

http://comptox.epa.gov/dashboard/

iCSS Dashboard (ToxCast data):

http://actor.epa.gov/dashboard/

**CPdat:** 

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

Underlying

Databases: **DSStox** (Distributed structure-searchable toxicity (DSSTox) public

database, Richard et al., 2002)

**ToxRefDB** (Animal Study data, Martin et al., 2009)

**CPCPdb** (Consumer Product Chemical Pathways databse,

Goldsmith et al, 2014)

R Packages: httk: High-Throughput Toxicokinetics (Pearce et al., in press)

https://cran.r-project.org/web/packages/httk/index.html

tcpl: ToxCast Data Analysis Pipeline (Filer et al., 2014)

https://cran.r-project.org/web/packages/tcpl/index.html



### **Conclusions**

- We would like to know more about the risk posed by thousands of chemicals in the environment which ones should we start with?
  - High throughput screening (HTS) provides a path forward for identifying potential hazard
  - Exposure and dosimetry provide real world context to hazards indicated by HTS
- Using in vitro methods developed for pharmaceuticals, we can relatively efficiently predict TK for large numbers of chemicals, but we are limited by analytical chemistry
- Using high throughput exposure approaches we can make coarse predictions of exposure
  - We are actively refining these predictions with new models and data
  - In some cases, upper confidence limit on current predictions is already many times lower than predicted hazard
- Expanded monitoring data (exposure surveillance) allows evaluation of model predictions
  - Are chemicals missing that we predicted would be there?
  - Are there unexpected chemicals?
- All data being made public:
  - R package "httk": https://CRAN.R-project.org/package=httk
  - The Chemistry Dashboard (A "Google" for chemicals) http://comptox.epa.gov/
  - Consumer Product Database: http://actor.epa.gov/cpcat/

52 of 54

Office of Research and Development



### References

- Bosgra, S., et al. "An improved model to predict physiologically based model parameters and their inter-individual variability from anthropometry." Critical reviews in toxicology 2012;42:751-767
- Dionisio, Kathie L., et al. "Exploring Consumer Exposure Pathways and Patterns of Use for Chemicals in the Environment." Toxicology Reports (2015)
- Egeghy, Peter P., et al. "The exposure data landscape for manufactured chemicals." Science of the Total Environment 414: 159-166 (2012)
- Filer, Dayne L.. "The ToxCast analysis pipeline: An R package for processing and modeling chemical screening data." US Environmental Protection Agency: http://www.epa. gov/ncct/toxcast/files/MySQL% 20Database/Pipeline\_Overview. pdf (2014)
- Goldsmith, M-R., et al. "Development of a consumer product ingredient database for chemical exposure screening and prioritization." Food and chemical toxicology 65 (2014): 269-279.
- Hahn, Otto, and Fritz Straßmann. "Über die Entstehung von Radiumisotopen aus Uran durch Bestrahlen mit schnellen und verlangsamten Neutronen." Naturwissenschaften 26.46 (1938): 755-756.
- Ingle, Brandall L., et al. "Informing the Human Plasma Protein Binding of Environmental Chemicals by Machine Learning in the Pharmaceutical Space: Applicability Domain and Limits of Predictability." Journal of Chemical Information and Modeling 56.11 (2016): 2243-2252.
- Isaacs, Kristin K., et al. "SHEDS-HT: An Integrated Probabilistic Exposure Model for Prioritizing Exposures to Chemicals with Near-Field and Dietary Sources." Environmental Science and Technology 48.21 (2014): 12750-12759.
- Isaacs, Kristin K., et al. "Characterization and prediction of chemical functions and weight fractions in consumer products." Toxicology Reports 3 (2016): 723-732.
- Jamei, et al. "The Simcyp® population-based ADME simulator."
   Expert opinion on drug metabolism & toxicology 2009b;5:211-223

- LaLone, Carlie A., et al. "Editor's Highlight: Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS): A Web-Based Tool for Addressing the Challenges of Cross-Species Extrapolation of Chemical Toxicity." Toxicological Sciences 153.2 (2016): 228-245.
- McNally, et al., "PopGen: a virtual human population generator." Toxicology 2014
- Newton, et al. "Novel Polyfluorinated Compounds Identified Using High Resolution Mass Spectrometry Downstream of Manufacturing Facilities near Decatur, Alabama," Envionmental Science and Technology 51(3): 1544-1552 (2017)
- O'Connell, Steven G., Laurel D. Kincl, and Kim A. Anderson.
   "Silicone wristbands as personal passive samplers."
   Environmental science & technology 48.6 (2014): 3327-3335.
- Park, Youngja, H., et al. "High-performance metabolic profiling of plasma from seven mammalian species for simultaneous environmental chemical surveillance and bioeffect monitoring." Toxicology 295:47-55 (2012)
- Pearce, Robert, et al. "httk: R Package for High-Throughput Toxicokinetics." Journal of Statistical Software, 20177
- Pearce, Robert, et al. "Evaluation and Calibration of High-Throughput Predictions of Chemical Distribution to Tissues." submitted.
- Phillips, Katherine A., et al. "High-throughput screening of chemicals as functional substitutes using structure-based classification models." Green Chemistry (2017).
- Phillips, Katherine A., et al. "Suspect Screening Analysis of Chemicals in Consumer Products", submitted.
- Price et al., "Instructions for Use of Software Physiological Parameters for PBPK Modeling Version 1.3 (P3MTM 1.3)." 2003
- Rager, Julia E., et al. "Linking high resolution mass spectrometry data with exposure and toxicity forecasts to advance highthroughput environmental monitoring." Environment International 88 (2016): 269-280.
- Rappaport, Stephen M., et al. "The blood exposome and its role in discovering causes of disease." Environmental Health Perspectives (Online) 122.8 (2014): 769.,

- Ring, Caroline, et al., "Identifying populations sensitive to environmental chemicals by simulating toxicokinetic variability", Environment International, 2017
- Ring, Caroline, et al., "Chemical Exposure Pathway Prediction for Screening and Priority-Setting", in preparation
- Schymanski, Emma L., et al. "Non-target screening with highresolution mass spectrometry: critical review using a collaborative trial on water analysis." Analytical and bioanalytical chemistry 407.21 (2015): 6237-6255.
- Sipes, Nisha, et al. "An Intuitive Approach for Predicting Potential Human Health Risk with the Tox21 10k Library", Environmental Science and Technology, in press
- Wallace et al., "The TEAM Study: Personal exposures to toxic substances in air, drinking water, and breath of 400 residents of New Jersey, North Carolina, and North Dakota." Environmental Research 43: 209-307 (1987)
- Wambaugh, John F., et al. "High-throughput models for exposure-based chemical prioritization in the ExpoCast project." Environmental science & technology 47.15 (2013): 8479-848.
- Wambaugh, John F., et al. "High Throughput Heuristics for Prioritizing Human Exposure to Environmental Chemicals." Environmental science & technology (2014).
- Wambaugh, John F., et al. "Toxicokinetic triage for environmental chemicals." Toxicological Sciences (2015): kfv118.
- Wetmore, Barbara A., et al. "Integration of dosimetry, exposure and high-throughput screening data in chemical toxicity assessment." Toxicological Sciences (2012): kfr254.
- Wetmore, Barbara A., et al. "Incorporating High-Throughput Exposure Predictions with Dosimetry-Adjusted In Vitro Bioactivity to Inform Chemical Toxicity Testing." Toxicological Sciences 148.1 (2015): 121-136.
- Zaldívar Comenges, José-Manuel, et al. "Modeling in vitro cellbased assays experiments: Cell population dynamics." Models of the Ecological Hierarchy: From Molecules to the Ecosphere 25 (2012): 51.