EPA
United States
Environmental Protection
Agency

Establishing Real World Context for High Throughput Toxicity Testing

John Wambaugh
National Center for Computational Toxicology
Office of Research and Development
U.S. Environmental Protection Agency

Presentation to Michigan State University
Institute for Integrative Toxicology
February 9, 2018

orcid.org/0000-0002-4024-534X

EPA Office of Research and Development

United States Environmental Protection Agency

- The Office of Research and Development (ORD) is the scientific research arm of EPA
  - 558 peer-reviewed journal articles in 2016
- Research is conducted by ORD's three national laboratories, four national centers, and two offices
  - Includes **National Center for Computational Toxicology** and **National Exposure Research Laboratory**
- 14 facilities across the country
- Six research programs
  - Includes **Chemical Safety for Sustainability**
- Research conducted by a combination of Federal scientists; contract researchers; and postdoctoral, graduate student, and post-baccalaureate trainees

ORD Facility in
Research Triangle Park, NC

Chemical Regulation in the United States

- Park *et al.* (2012): At least 3221 chemicals in pooled human blood samples, many appear to be exogenous
- A tapestry of laws covers the chemicals people are exposed to in the United States (Breyer, 2009)
- Different testing requirements exist for food additives, pharmaceuticals, and pesticide active ingredients (NRC, 2007)

November 29, 2014

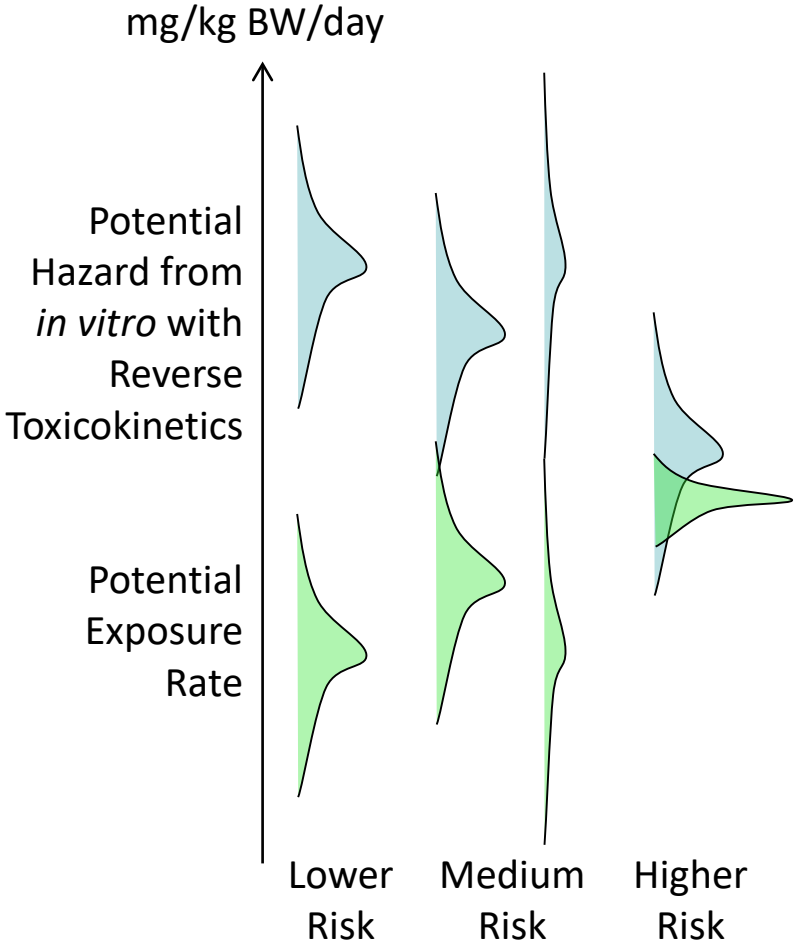
Chemical Regulation in the United States

- Most other chemicals, ranging from industrial waste to dyes to packing materials, are covered by the Toxic Substances Control Act (TSCA)
  - Thousands of chemicals on the market were either "grandfathered" in or were allowed without experimental assessment of hazard, toxicokinetics, or exposure
  - Thousands of new chemical use submissions are made to the EPA every year
- TSCA was updated in June, 2016 to allow evaluation of these and other chemicals
  - Methods are being developed to prioritize these existing and new chemicals for testing

November 29, 2014

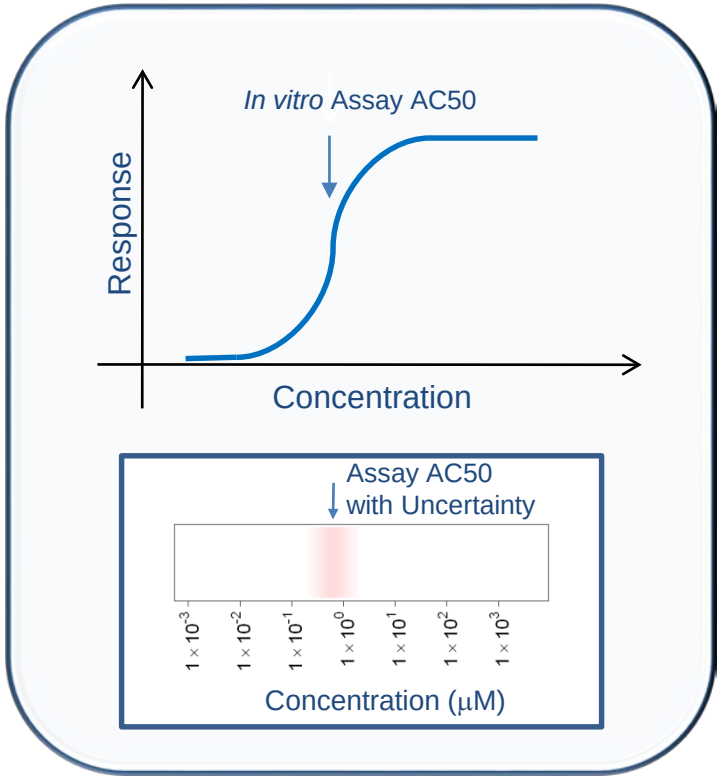
Chemical Risk = Hazard + Exposure

- National Research Council (1983) identified chemical risk as a function of both inherent hazard and exposure
- To address thousands of chemicals, we need to use “high throughput methods” to prioritize those chemicals most worthy of additional study
- **High throughput risk prioritization** needs:
  1. high throughput **hazard** characterization (from HTT project)
  2. high throughput **exposure** forecasts
  3. high throughput **toxicokinetics** (*i.e.*, dosimetry) linking hazard and exposure



High-Throughput Screening

- We might estimate points of departure (concentrations causing relevant bioactivity) *in vitro* using high throughput screening (HTS)
- **Tox21**: Examining >8,000 chemicals using ~50 assays intended to identify interactions with biological pathways (Schmidt, 2009)
- **ToxCast**: For a subset (>2000) of Tox21 chemicals ran >1100 additional assays (Kavlock *et al.*, 2012)
- Most assays conducted in dose-response format (identify 50% activity concentration – AC50 – and efficacy if data described by a Hill function, Filer *et al.*, 2016)
- All data is public: http://comptox.epa.gov/dashboard/



Risk Assessment in the 21st Century

United States Environmental Protection Agency

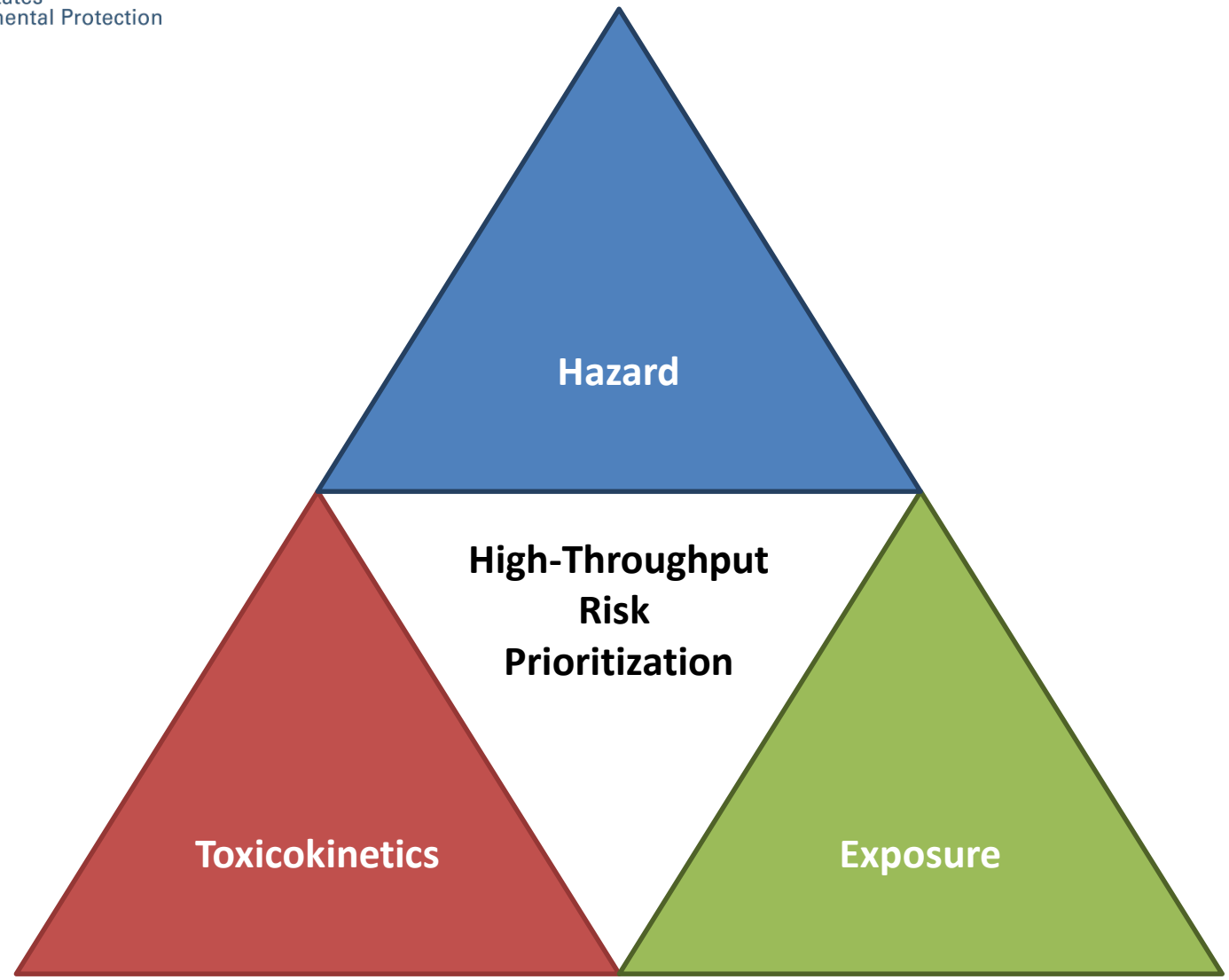
"Translation of high-throughput data into risk-based rankings is an important application of exposure data for chemical priority-setting. Recent advances in high-throughput toxicity assessment, notably the ToxCast and Tox21 programs (see Chapter 1), and in high-throughput computational exposure assessment (Wambaugh et al. 2013, 2014) have enabled first-tier risk-based rankings of chemicals on the basis of margins of exposure…"

"…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**."

January 5, 2017

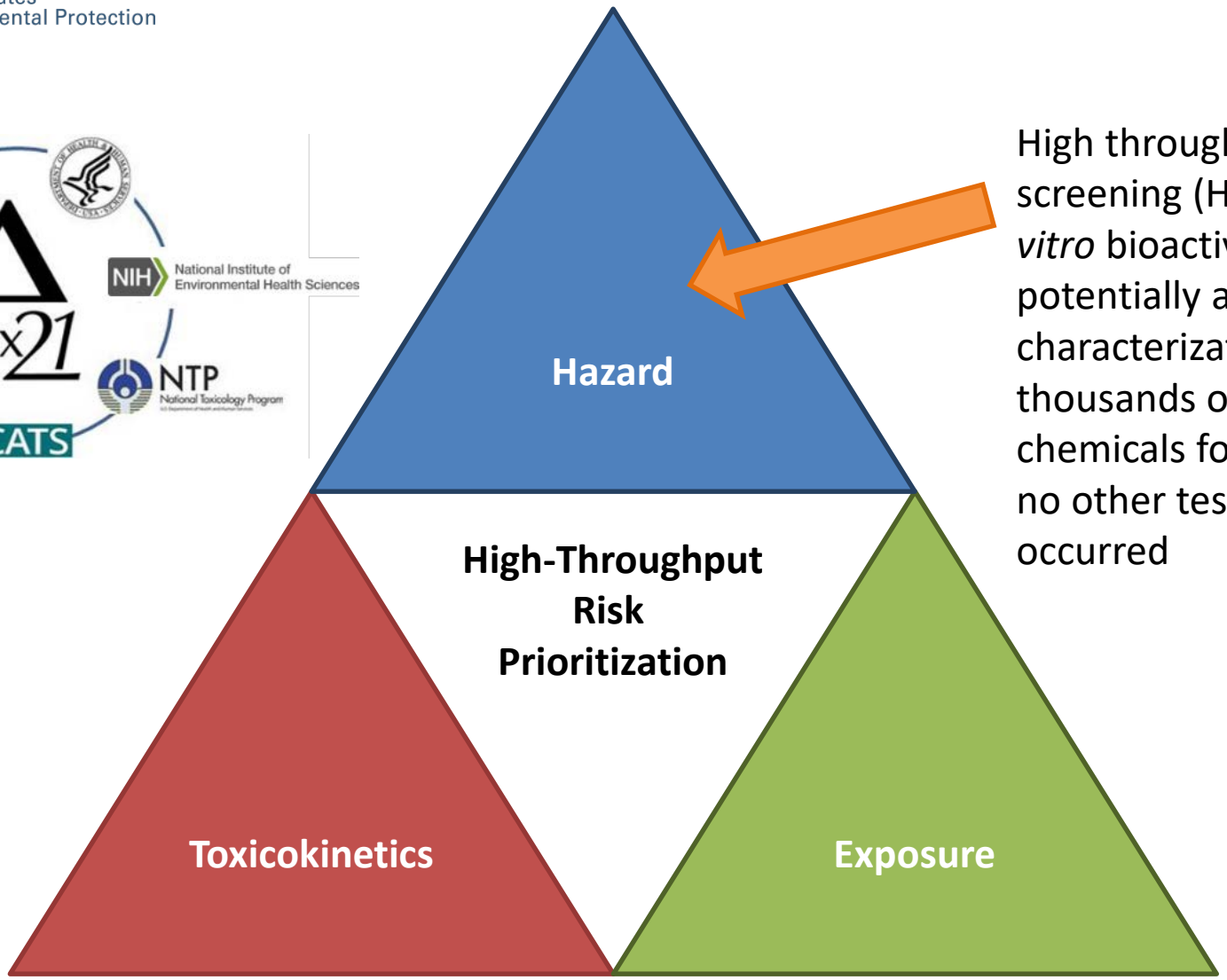
Three Components for Chemical Risk

United States Environmental Protection Agency



High-Throughput Risk Prioritization

EPA United States Environmental Protection Agency



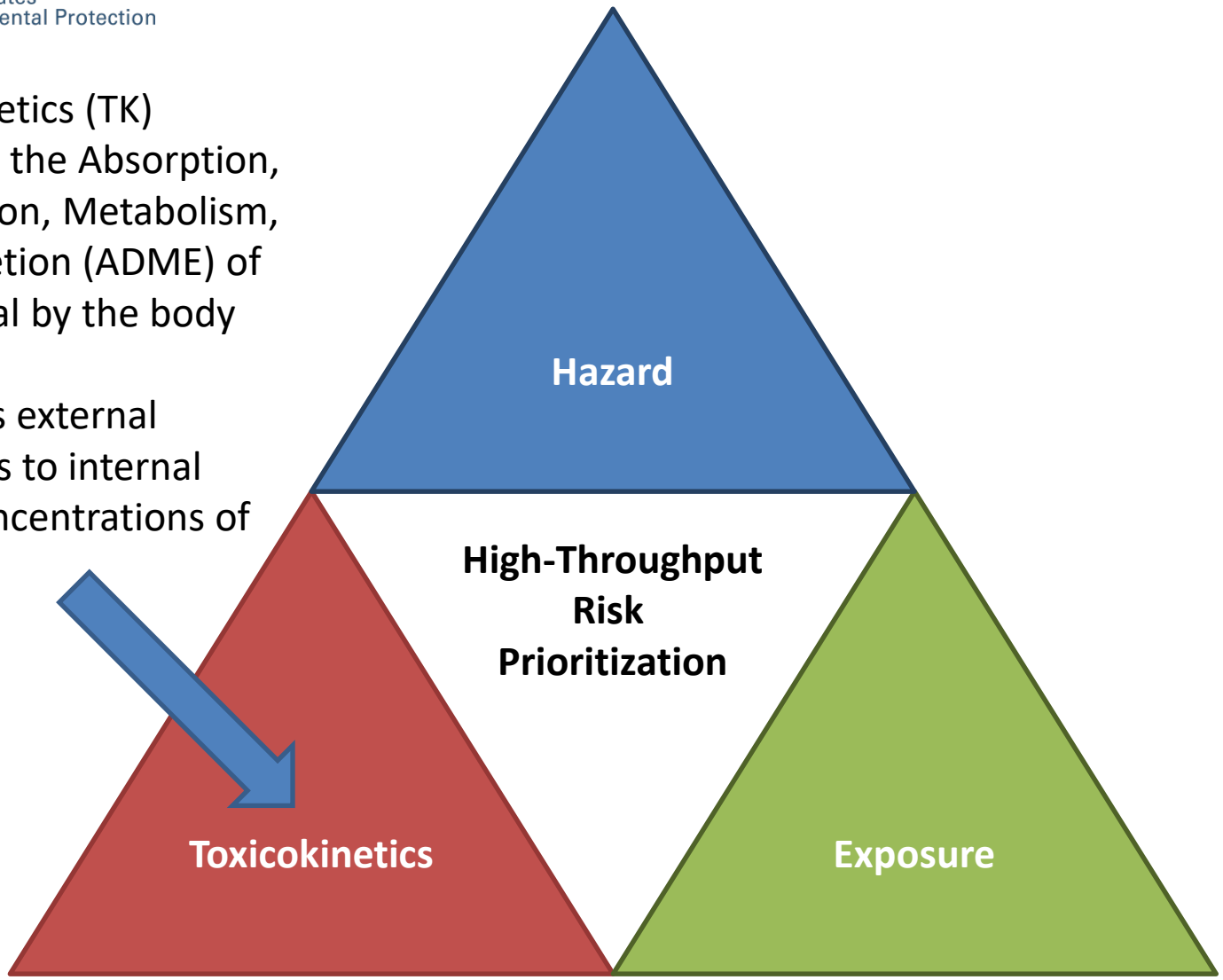
High throughput screening (HTS) for *in vitro* bioactivity potentially allows characterization of thousands of chemicals for which no other testing has occurred

High Throughput Toxicokinetics (HTTK)

United States Environmental Protection Agency

Toxicokinetics (TK) describes the Absorption, Distribution, Metabolism, and Excretion (ADME) of a chemical by the body

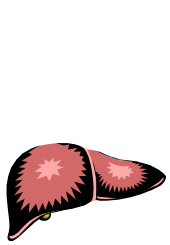
TK relates external exposures to internal tissue concentrations of chemical

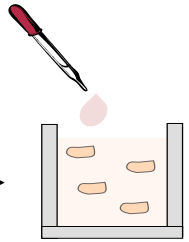


High-Throughput Toxicokinetics (HTTK)

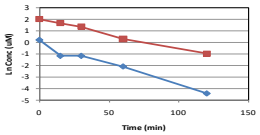
United States Environmental Protection Agency

- **Most chemicals do not have TK data** – we use *in vitro* HTTK methods adapted from pharma to fill gaps
- In drug development, HTTK methods estimate therapeutic doses for clinical studies – predicted concentrations are typically on the order of values measured in clinical trials (Wang, 2010)





Human Hepatocytes (10 donor pool)



Intrinsic Hepatic Clearance (Cl_{int})

Measurements require chemical-specific methods for concentration

In Vitro - In Vivo Extrapolation

Predicted Plasma Concentrations

Human Plasma (6 donor pool)

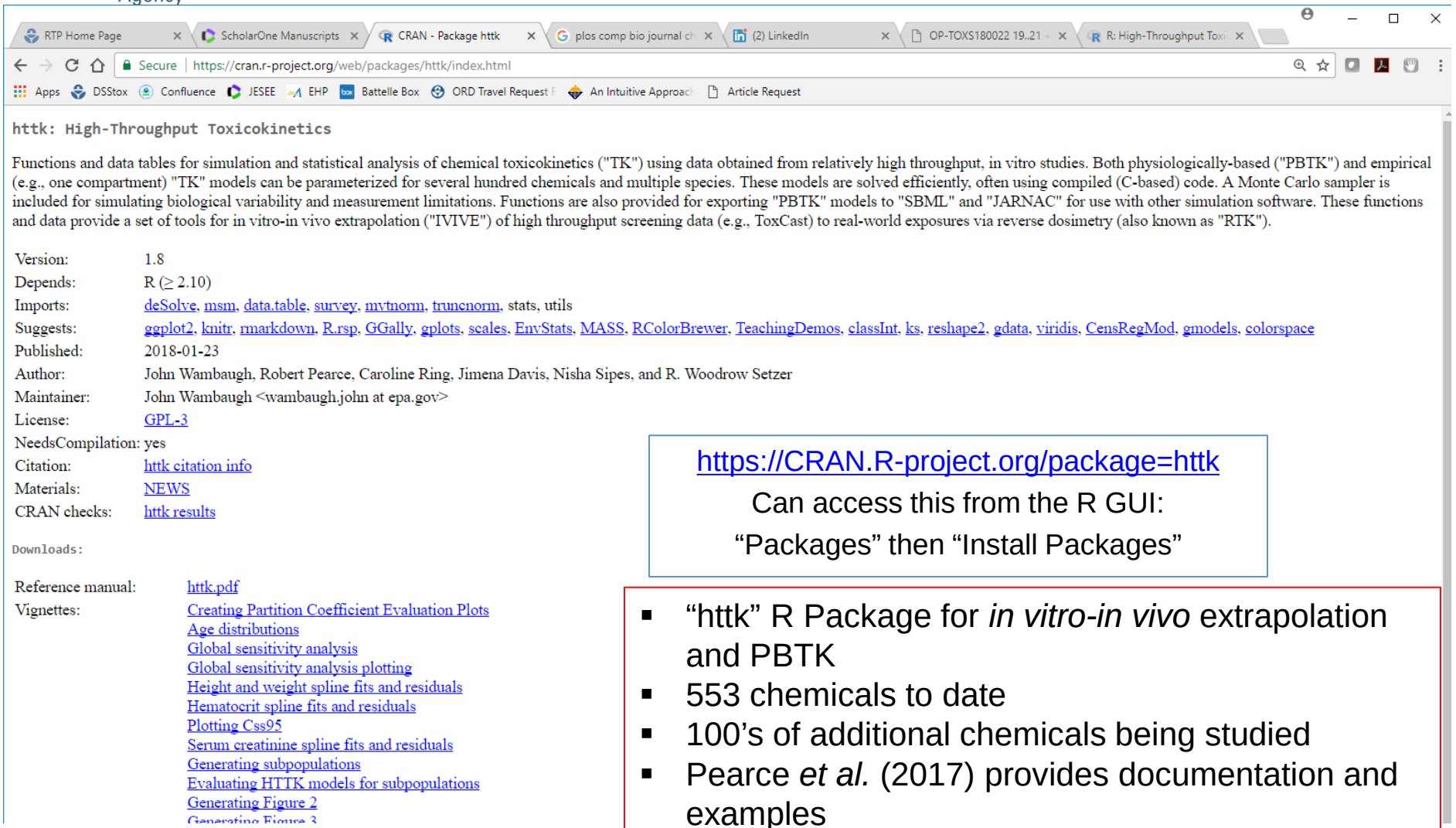
Plasma Protein Binding (f_{up})

Rotroff et al. (2010) 35 chemicals
Wetmore et al. (2012) +204 chemicals
Wetmore et al. (2015) +163 chemicals

Figure from Barbara Wetmore

Open Source Tools and Data for HTTK

United States Environmental Protection Agency



RTP Home Page | ScholarOne Manuscripts | CRAN - Package httk | plos comp bio journal | (2) LinkedIn | OP-TOXS180022 19..21 | R: High-Throughput Toxi

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

Apps DSStox Confluence JESEE EHP Battelle Box ORD Travel Request An Intuitive Approac Article Request

httk: High-Throughput Toxicokinetics

Functions and data tables for simulation and statistical analysis of chemical toxicokinetics ("TK") using data obtained from relatively high throughput, in vitro studies. Both physiologically-based ("PBTK") and empirical (e.g., one compartment) "TK" models can be parameterized for several hundred chemicals and multiple species. These models are solved efficiently, often using compiled (C-based) code. A Monte Carlo sampler is included for simulating biological variability and measurement limitations. Functions are also provided for exporting "PBTK" models to "SBML" and "JARNAC" for use with other simulation software. These functions and data provide a set of tools for in vitro-in vivo extrapolation ("IVIVE") of high throughput screening data (e.g., ToxCast) to real-world exposures via reverse dosimetry (also known as "RTK").

Version:	1.8
Depends:	R (≥ 2.10)
Imports:	deSolve, msm, data.table, survey, mvtnorm, truncnorm, stats, utils
Suggests:	ggplot2, knitr, rmarkdown, R.rsp, GGally, gplots, scales, EnvStats, MASS, RColorBrewer, TeachingDemos, classInt, ks, reshape2, gdata, viridis, CensRegMod, gmodels, colorspace
Published:	2018-01-23
Author:	John Wambaugh, Robert Pearce, Caroline Ring, Jimena Davis, Nisha Sipes, and R. Woodrow Setzer
Maintainer:	John Wambaugh <wambaugh.john at epa.gov>
License:	GPL-3
NeedsCompilation:	yes
Citation:	httk citation info
Materials:	NEWS
CRAN checks:	httk results

Downloads:

Reference manual:	httk.pdf
Vignettes:	Creating Partition Coefficient Evaluation Plots 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

https://CRAN.R-project.org/package=httk

Can access this from the R GUI:

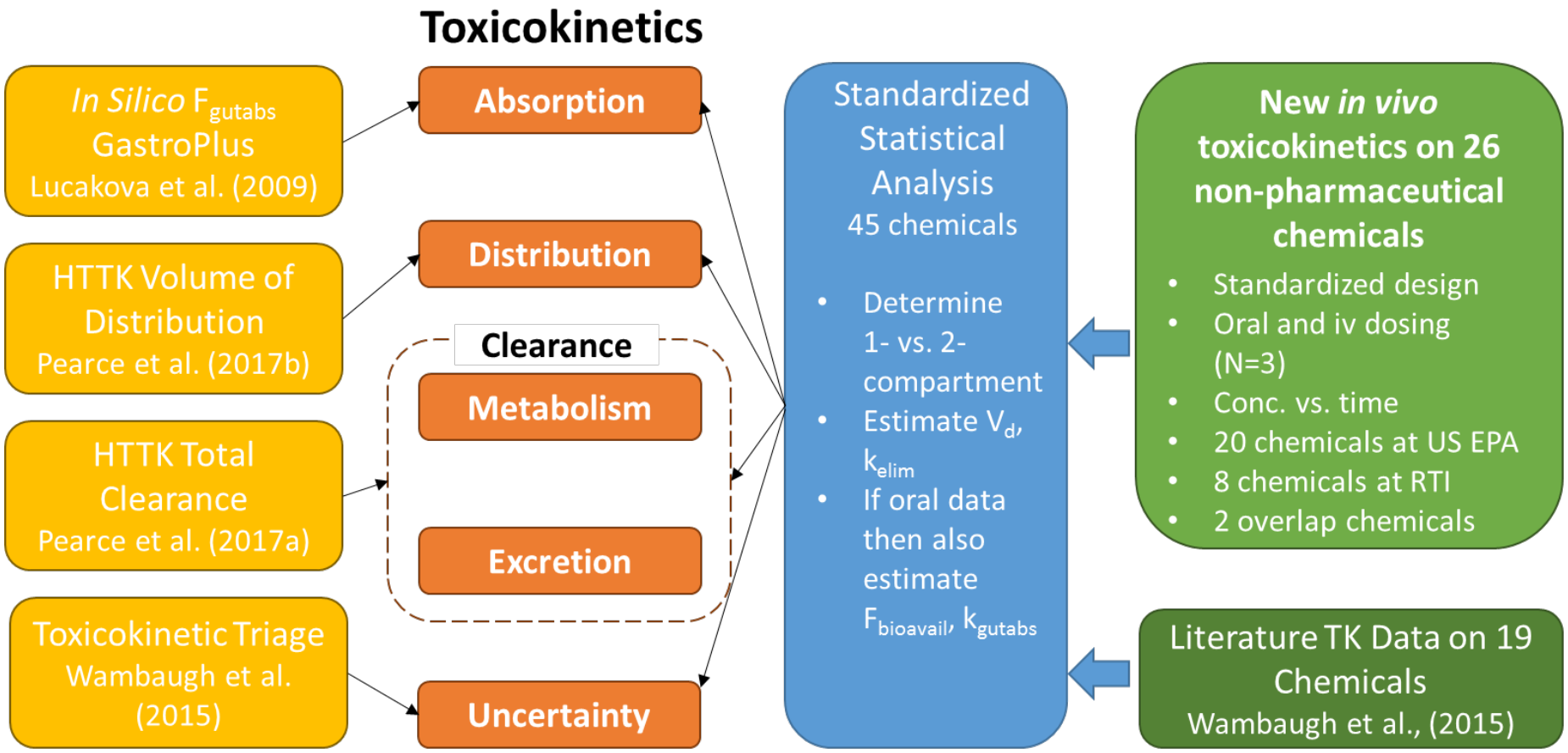
"Packages" then "Install Packages"

- "httk" R Package for *in vitro-in vivo* extrapolation 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

Building Confidence in HTTK

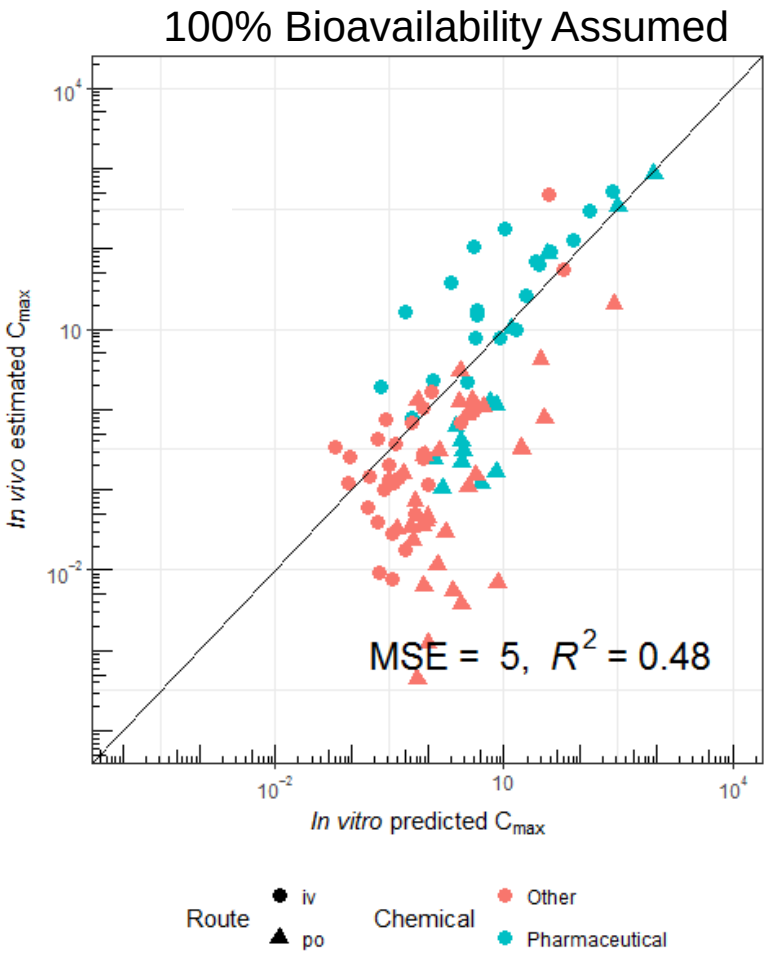
We collected new data for 26 chemicals more commonly associated with non-therapeutic and/or unintentional exposure

Minimal design – six animals per study (3 dosed per oral / 3 iv)

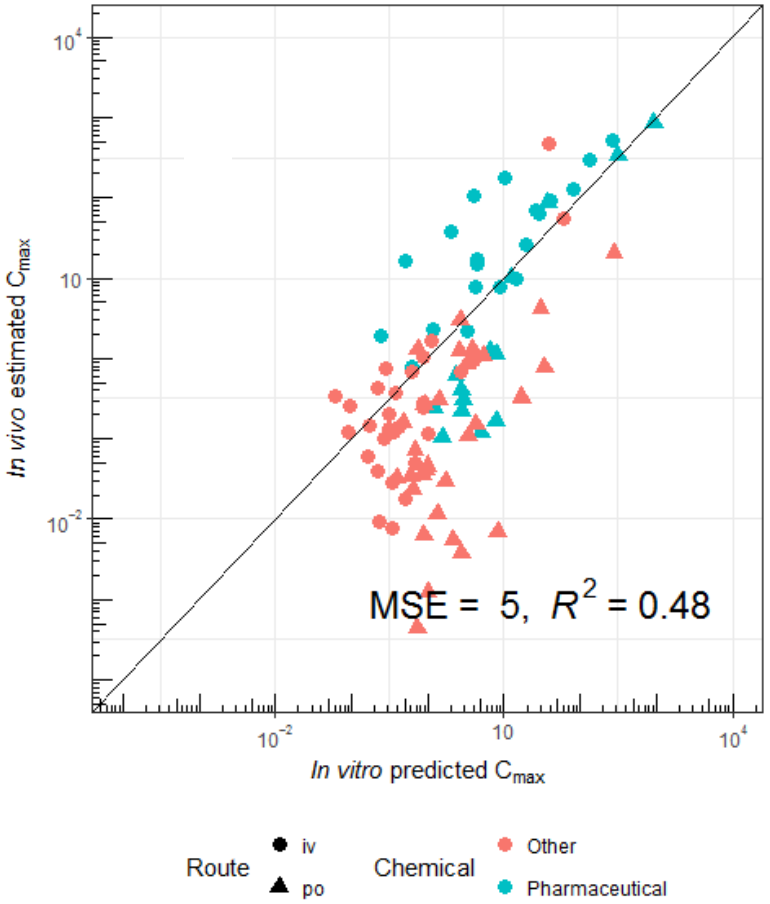


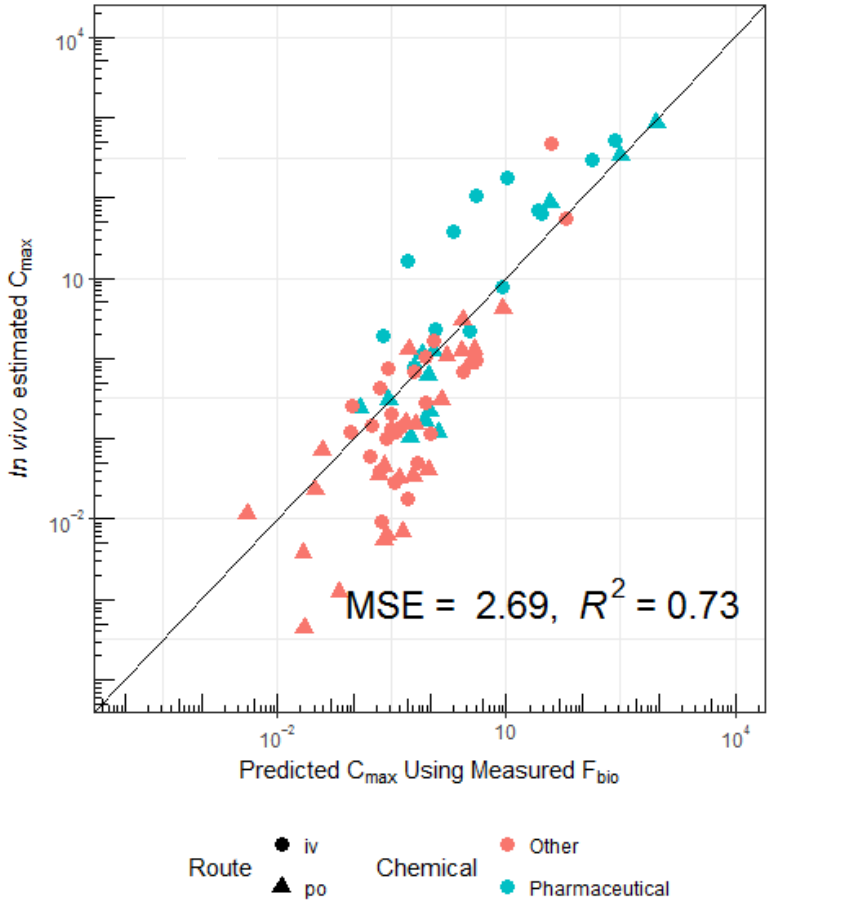
Wambaugh et al. (Tox. Sci., *just accepted*)

Evaluating HTTK



Evaluating HTTK





Greg Honda (NCCT) will give a SOT2018 presentation on using Caco2 *in vitro* data to predict absorption for ~300 ToxCast chemicals

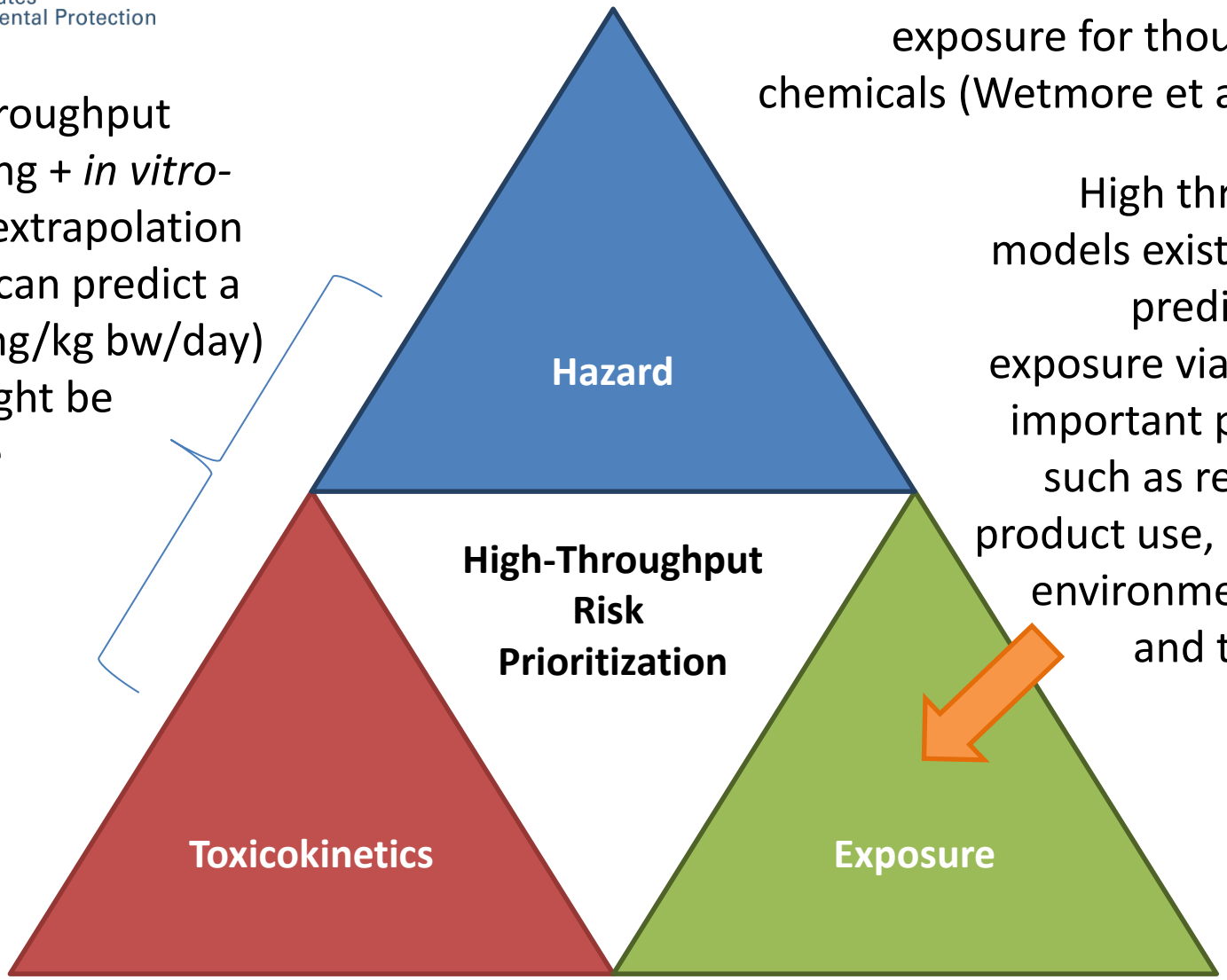
New Exposure Data and Models

EPA United States Environmental Protection Agency

Need methods to forecast exposure for thousands of chemicals (Wetmore et al., 2015)

High throughput screening + *in vitro-in vivo* extrapolation (IVIVE) can predict a dose (mg/kg bw/day) that might be adverse

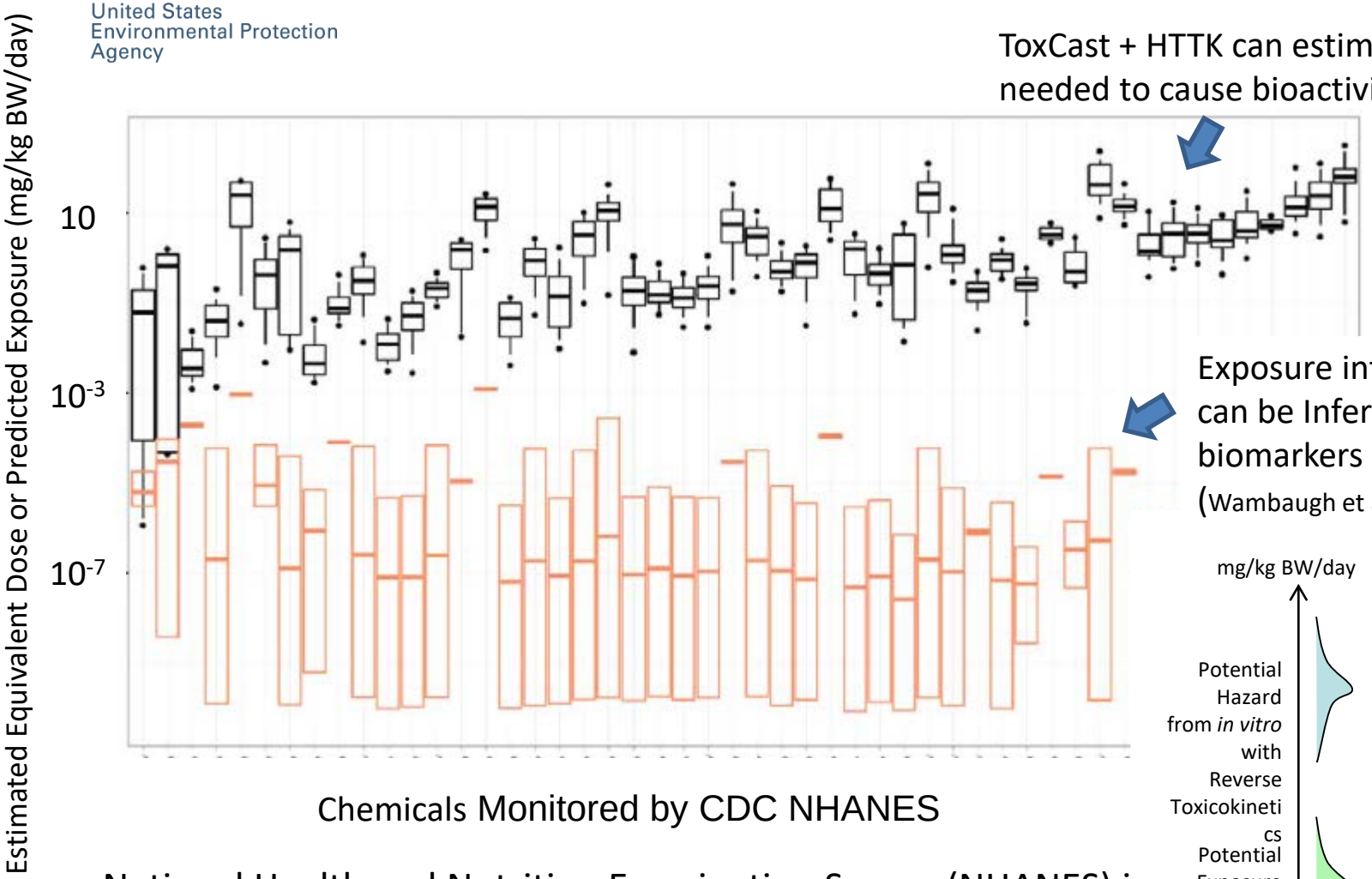
High throughput models exist to make predictions of exposure via specific, important pathways such as residential product use, diet, and environmental fate and transport



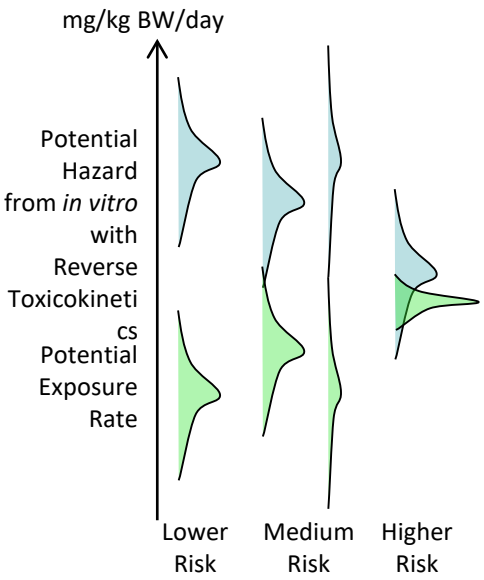
High Throughput Risk Prioritization

United States Environmental Protection Agency

ToxCast + HTTK can estimate doses needed to cause bioactivity



Exposure intake rates can be Inferred from biomarkers (Wambaugh et al., 2014)

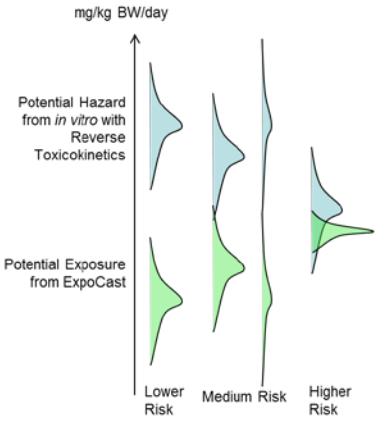


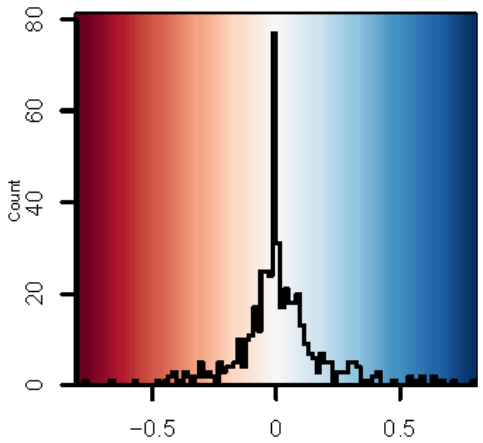
National Health and Nutrition Examination Survey (NHANES) is an ongoing survey that covers ~10,000 people every two years

Ring *et al*. (2017)

Life-stage and Demographic Specific Predictions

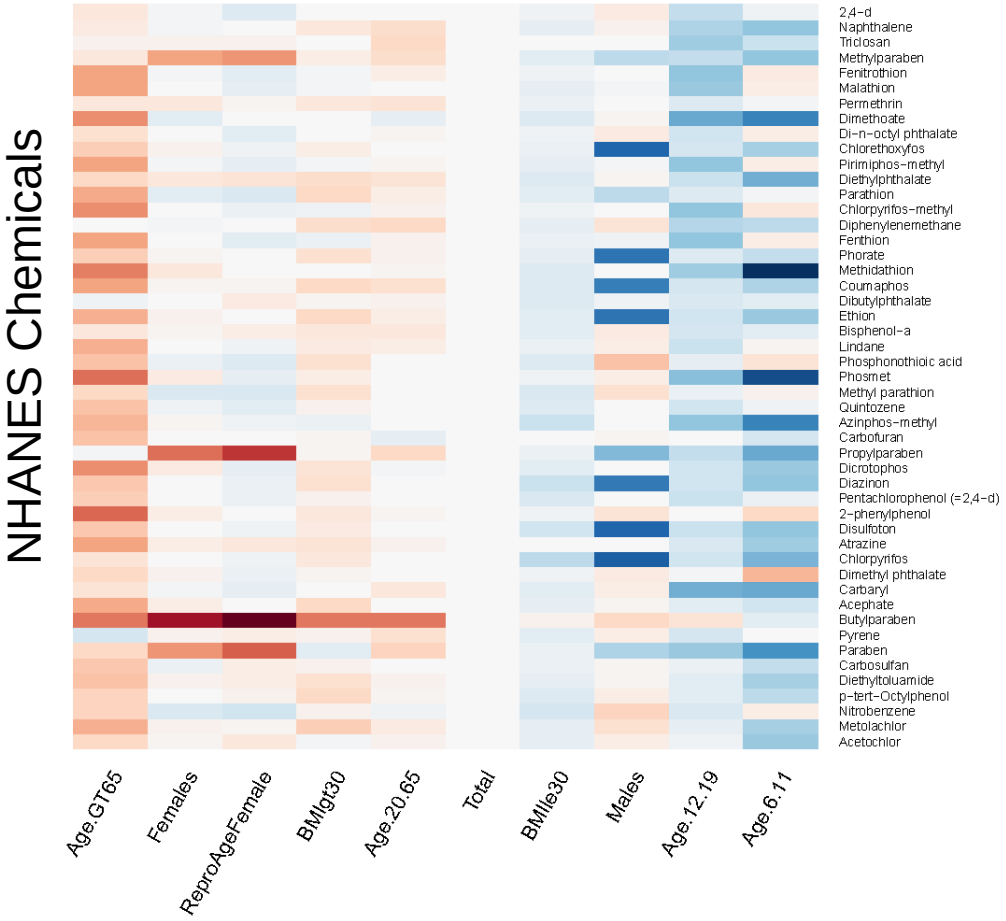
- Can calculate margin between bioactivity and exposure for specific populations





Change in Risk Relative to Total Population

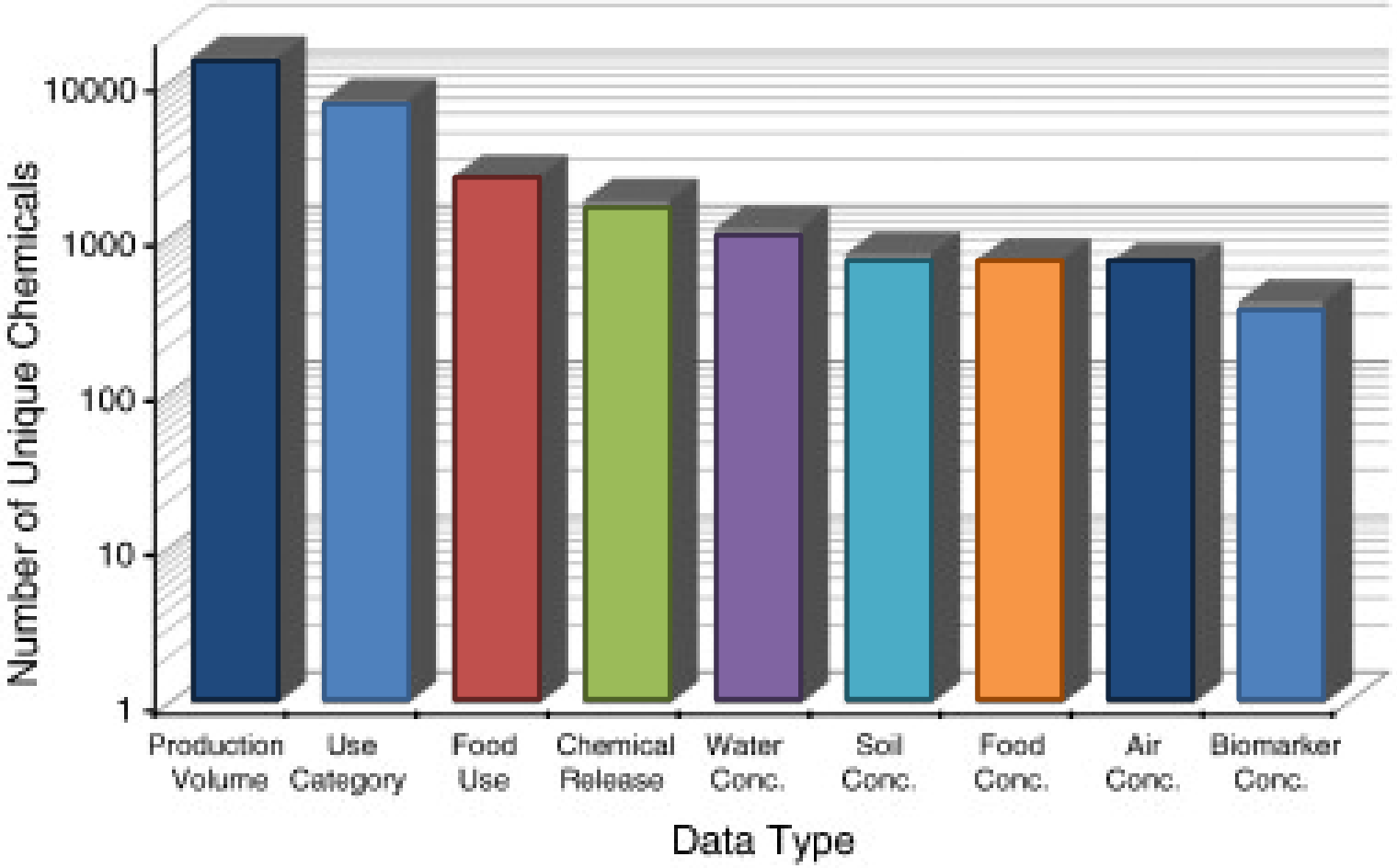
Change in Activity:Exposure Ratio



Ring *et al*. (2017)

Limited Available Data for Exposure Estimation

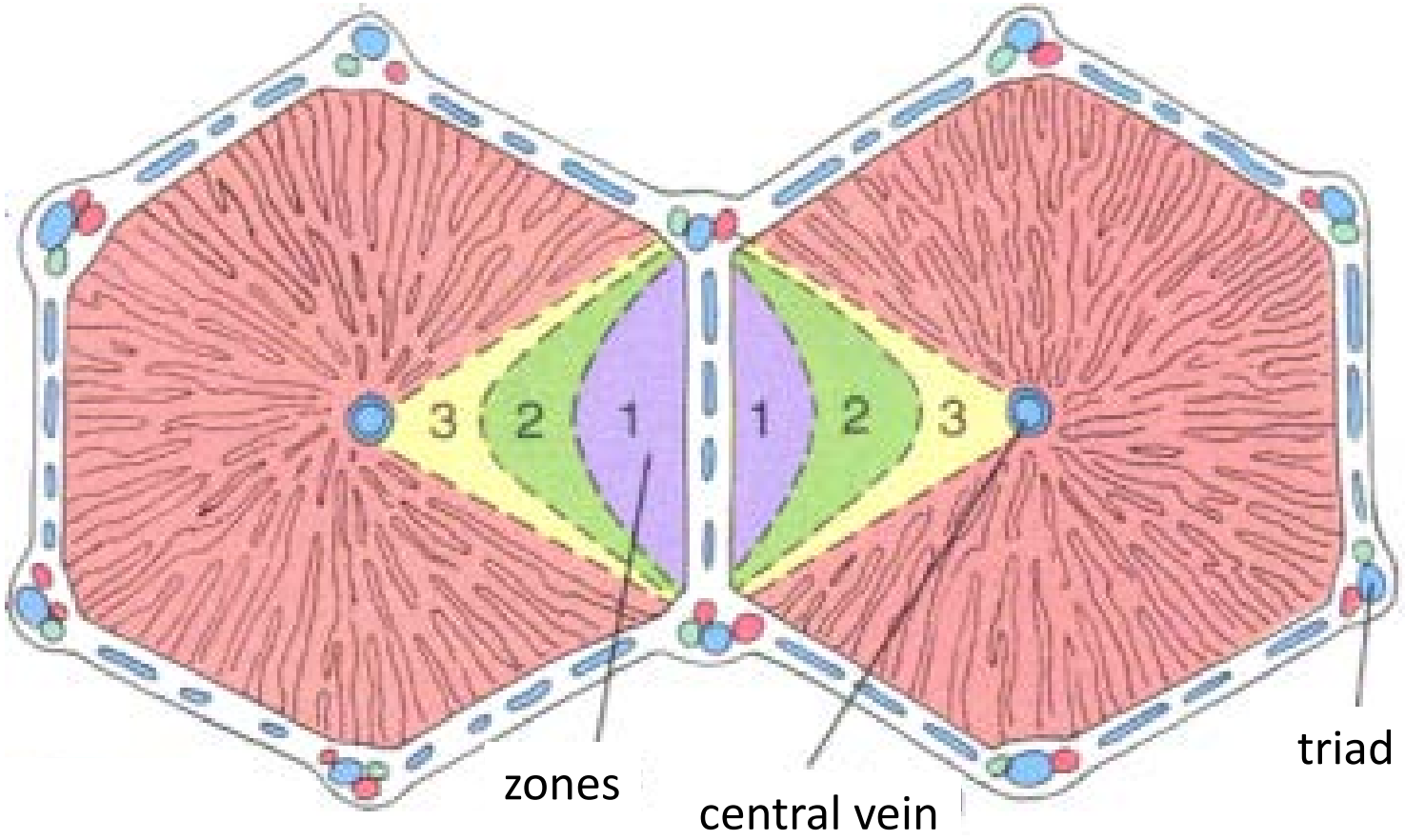
Most chemicals lack public exposure-related data beyond production volume (Egeghy et al., 2012)



Can we use models to generate the exposure information we need?

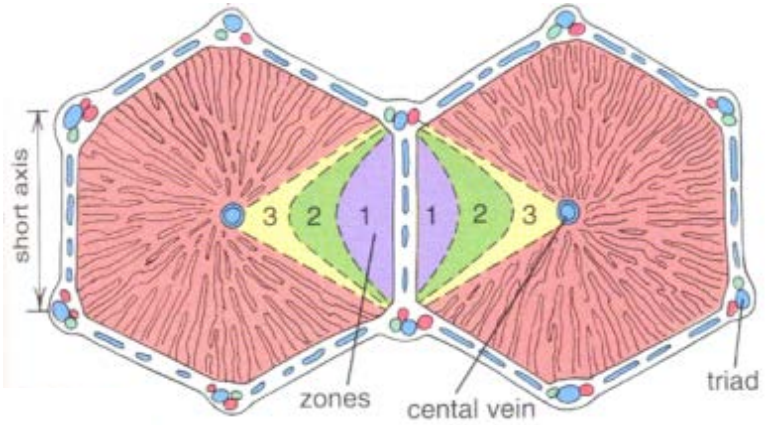
Computational Approaches: Modeling

The liver is composed of hepatic lobules



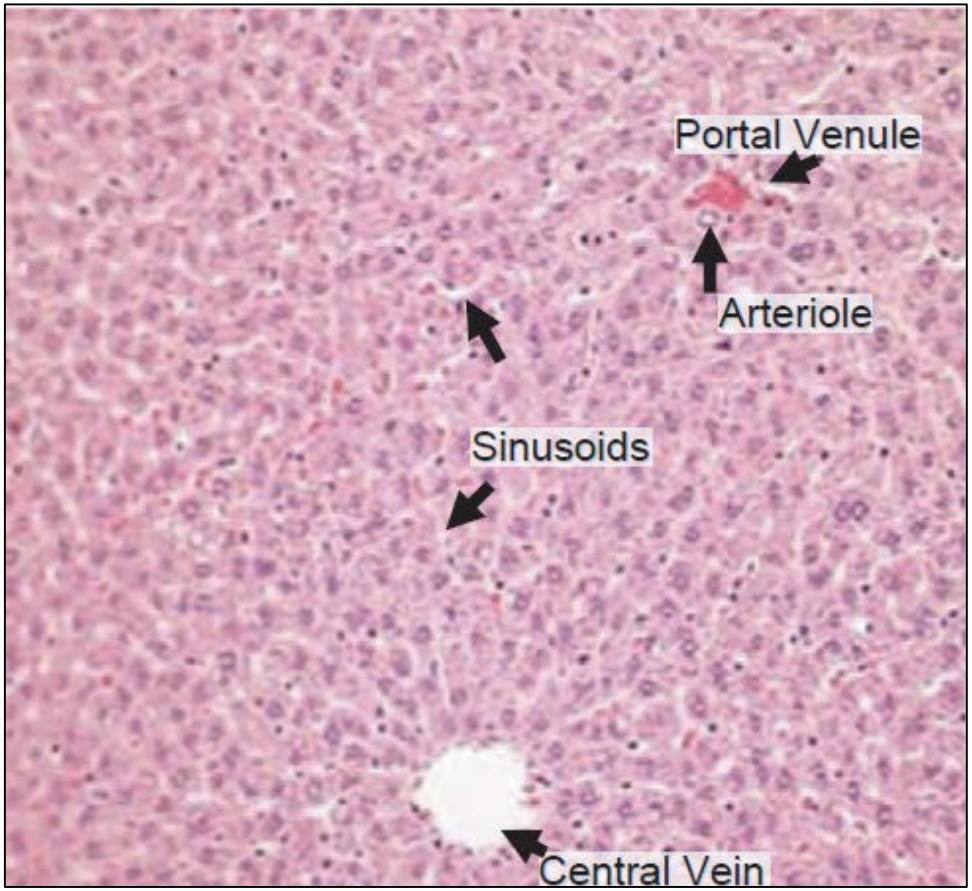
University of South Dakota

When Models Meet Real Biological Variability

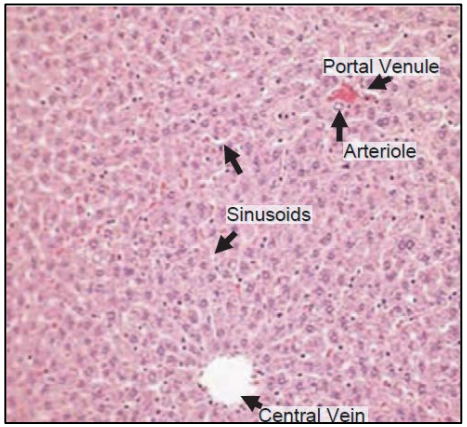


- Actual lobules are much messier (variable) (Crawford, et al., 1988)
- Further, pathology calls involve subjectivity
- You need to understand both the system being modeled and the data generation process

Rockett et al. (2006)



Pattern Recognition



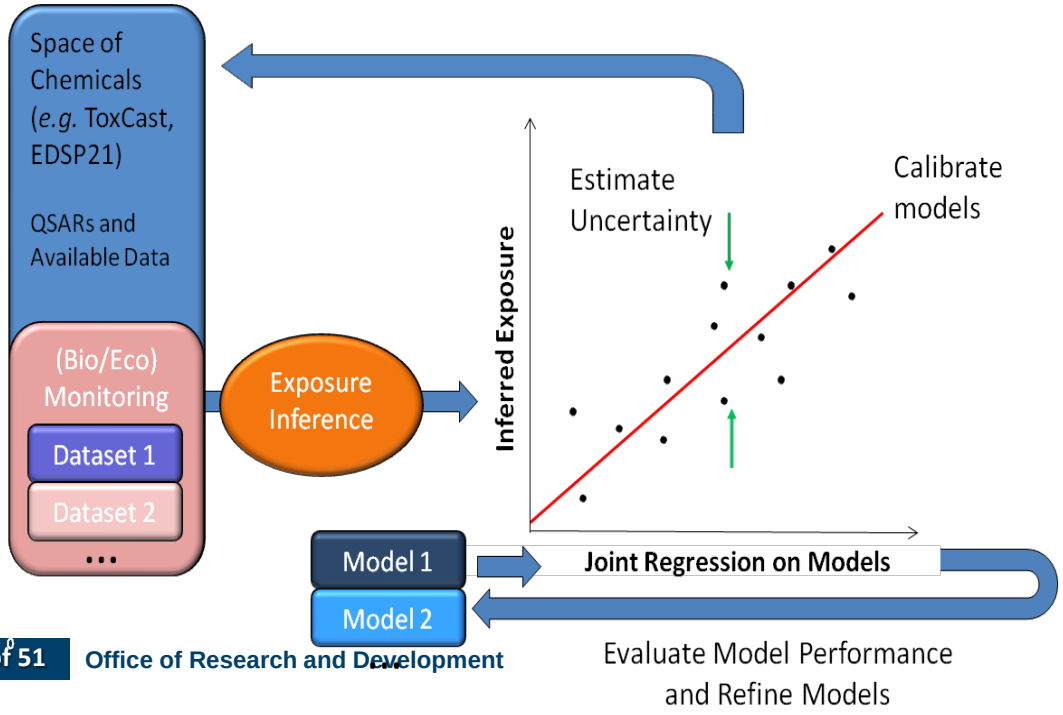
Rockett et al. (2006)

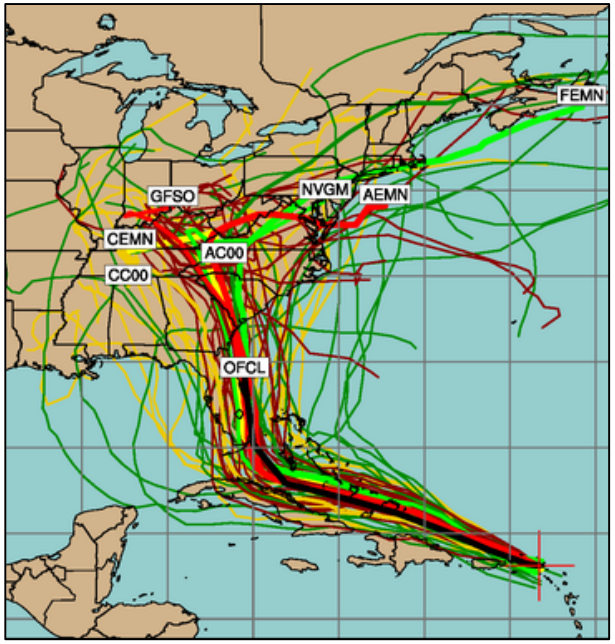
Teatra Sociale, Como, Italy

- The underlying rules of system
- Each hepatocyte needs to get oxygen, state depends on degree of hypoxia, endogenous chemical signaling, and history of exposure to exogenous chemicals

Consensus Exposure Predictions with the SEEM Framework

- Different exposure models incorporate **knowledge**, **assumptions**, and **data** (Macleod, et al., 2010)
- We incorporate multiple models (including SHEDS-HT, ExpoDat) into consensus predictions for 1000s of chemicals within the **Systematic Empirical Evaluation of Models (SEEM) framework**
- Evaluation is similar to a sensitivity analysis: What models are working? What data are most needed?





Integrating Multiple Models

Collaboration on High Throughput Exposure Predictions

Jon Arnot, Deborah H. Bennett, Peter P. Egeghy, Peter Fantke, Lei Huang, Kristin K. Isaacs, Olivier Jolliet, Hyeong-Moo Shin, Katherine A. Phillips, Caroline Ring, R. Woodrow Setzer, John F. Wambaugh, Johnny Westgate

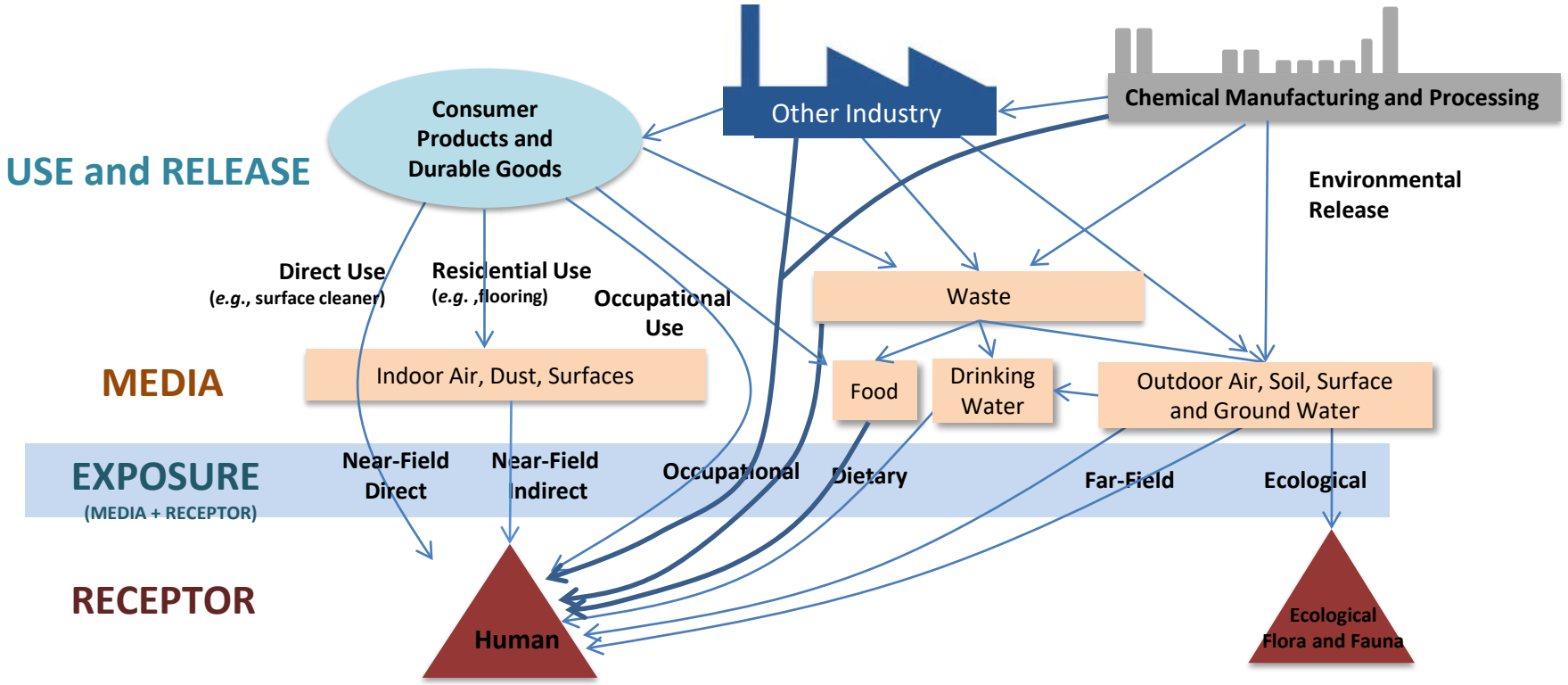
Model	Version	Reference	Dietary	Near-Field	Far-Field Pesticide	Far-Field Industrial
EPA Stochastic Human Exposure Dose Simulator High Throughput (SHEDS-HT) Near-Field Direct	2017	Isaacs, et al. (2014)		1119		
SHEDS-HT Near-field Indirect	2017	Isaacs, et al. (2014)		645		
Shin-Bennett	2017	Shin et al. , in preparation		1221		
Food Contact Substance Migration Model	2017	Biryol et al. (2017)	940			
EPA Pesticide Reregistration Eligibility Documents (REDs) Exposure Assessments	2015	Wetmore et al. (2012, 2015)			239	
Risk Assessment IDentification And Ranking (RAIDAR) Far-Field	2.941	Arnot et al. (2006)			7511	7511
RAIDAR-ICE Near-Field	0.803	Arnot et al., in preparation		615		
United Nations Environment Program and Society for Environmental Toxicology and Chemistry toxicity model (USETox) Pesticide Scenario	1.01	Rosenbaum (2008)			790	
USEtox Industrial Scenario	1.01	Rosenbaum (2008)				7184
EPA Inventory Update Reporting and Chemical Data Reporting	2015	US EPA (2018)	7856	7856	7856	7856
FDA Cumulative Estimated Daily Intake (CDI)	2017	US FDA (2017)	748			
Stockholm Convention of Banned Persistent Organic Pollutants	2017	2017			22	225

Collaboration on High Throughput Exposure Predictions

Jon Arnot, Deborah H. Bennett, Peter P. Egeghy, Peter Fantke, Lei Huang, Kristin K. Isaacs, Olivier Jolliet, Hyeong-Moo Shin, Katherine A. Phillips, Caroline Ring, R. Woodrow Setzer, John F. Wambaugh, Johnny Westgate

el	Version	Reference	Dietary	Near-Field	Far-Field Pesticide	Far-Field Industrial
Stochastic Human Exposure Dose Simulator High ughput (SHEDS-HT) Near-Field Direct	2017	Isaacs, et al. (2014)		1119		
DS-HT Near-field Indirect	2017	Isaacs, et al. (2014)		645		
-Bennett	2017	Shin et al. , in preparation		1221		
Contact Substance Migration Model	2017	Biryol et al. (2017)	940			
EPA Pesticide Reregistration Eligibility Documents (REDs) Exposure Assessments	2015	Wetmore et al. (2012, 2015)			239	
Risk Assessment IDentification And Ranking (RAIDAR) Far-Field	2.941	Arnot et al. (2006)			7511	7511
RAIDAR-ICE Near-Field	0.803	Arnot et al., in preparation		615		
United Nations Environment Program and Society for Environmental Toxicology and Chemistry toxicity model (USETox) Pesticide Scenario	1.01	Rosenbaum (2008)			790	
USEtox Industrial Scenario	1.01	Rosenbaum (2008)				7184
EPA Inventory Update Reporting and Chemical Data Reporting	2015	US EPA (2018)	7856	7856	7856	7856
FDA Cumulative Estimated Daily Intake (CDI)	2017	US FDA (2017)	748			
Stockholm Convention of Banned Persistent Organic Pollutants	2017	2017			22	225

Forecasting Exposure is a Systems Problem



- **Exposure event unobservable:** Can try to predict exposure by characterizing pathway
- Some pathways have much higher average exposures: In home "Near field" sources significant (Wallace, *et al.,* 1987)

Figure from Kristin Isaacs

Knowledge of Exposure Pathways Limits High Throughput Exposure Models

"In particular, the assumption that 100% of [quantity emitted, applied, or ingested] is being applied to each individual use scenario is a very conservative assumption for many compound / use scenario pairs."

Environmental Science & Technology

Article

pubs.acs.org/est

Risk-Based High-Throughput Chemical Screening and Prioritization using Exposure Models and in Vitro Bioactivity Assays

Hyeong-Moo Shin,[*,†] Alexi Ernstoff,[‡,§] Jon A. Arnot,[∥,⊥,#] Barbara A. Wetmore,[▽] Susan A. Csiszar,[§] Peter Fantke,[‡] Xianming Zhang,[○] Thomas E. McKone,[◆,¶] Olivier Jolliet,[§] and Deborah H. Bennett[†]

[†]Department of Public Health Sciences, University of California, Davis, California 95616, United States

[‡]Quantitative Sustainability Assessment Division, Department of Management Engineering, Technical University of Denmark, Kgs. Lyngby 2800, Denmark

[§]Department of Environmental Health Sciences, University of Michigan, Ann Arbor, Michigan 48109, United States

[∥]ARC Arnot Research and Consulting, Toronto, Ontario M4M 1W4 , Canada

[⊥]Department of Physical and Environmental Sciences, University of Toronto, Scarborough, Toronto, Ontario M1C 1A4, Canada

[#]Department of Pharmacology and Toxicology, University of Toronto, Toronto, Ontario M5S 1A8, Canada

[▽]The Hamner Institutes for Health Sciences, Research Triangle Park, North Carolina 27709, United States

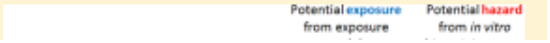
[○]Harvard School of Public Health and School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, United States

[◆]Environmental Energy Technologies Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720 , United States

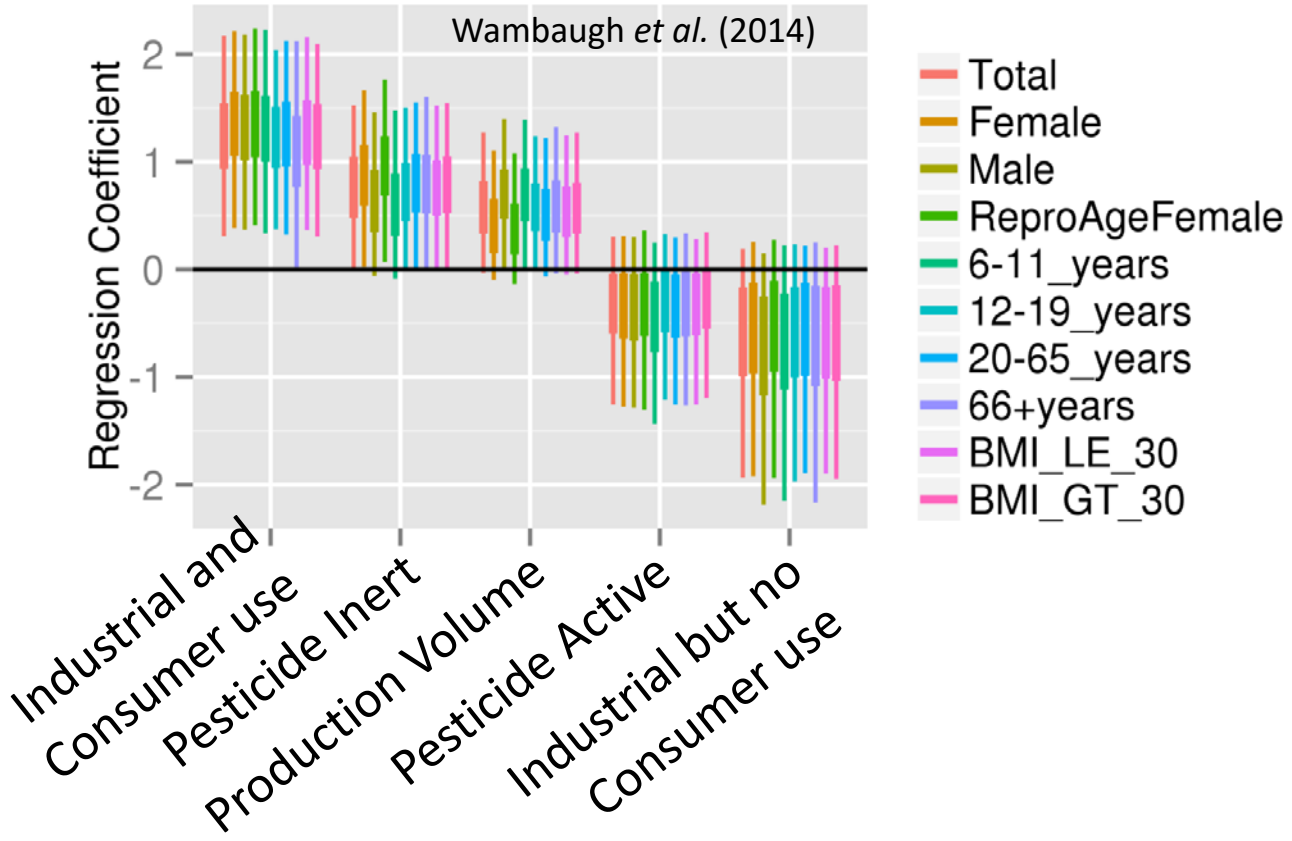
[¶]School of Public Health, University of California, Berkeley, California 94720, United States

Supporting Information

ABSTRACT: We present a risk-based high-throughput screening



Heuristics of Exposure



- 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
- Chemical use identifies relevant pathways
- Some pathways have much higher average exposures (Wallace et al., 1987)

Chemical Use: Chemicals and Products Database

United States Environmental Protection Agency

MSDS Data

Occurrence and quantitative chemical composition

Broad "index" of chemical uses

CPCat

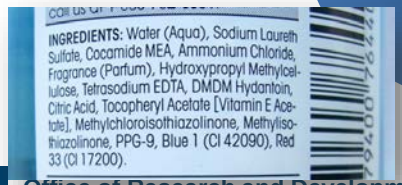
CPDat

EPA United States Environmental Protection Agency
Chemistry Dashboard

Functional Use Data

Ingredient Lists

Occurrence data



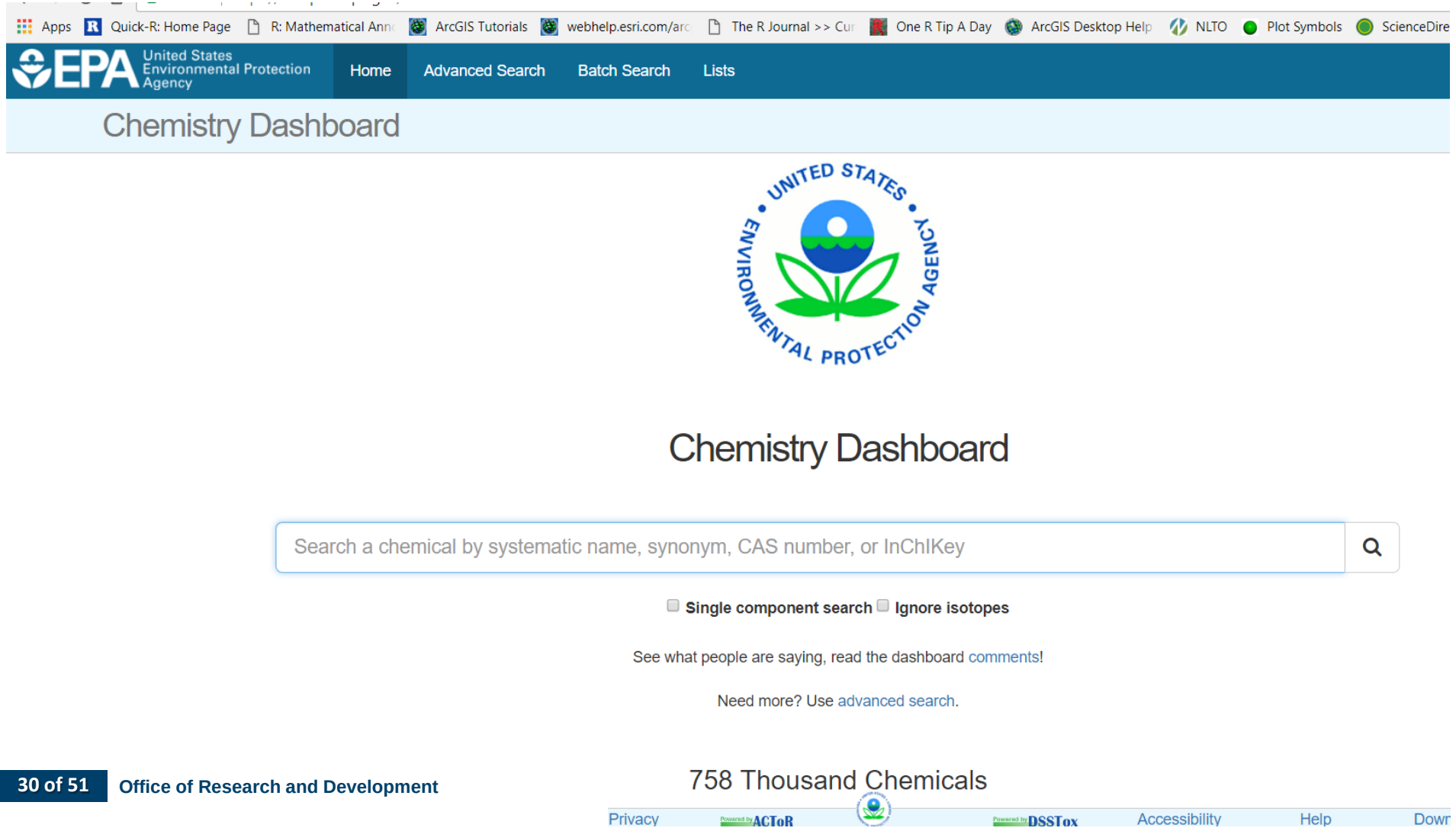
Measured Data

Targeted and non-targeted measurement of chemicals in consumer products

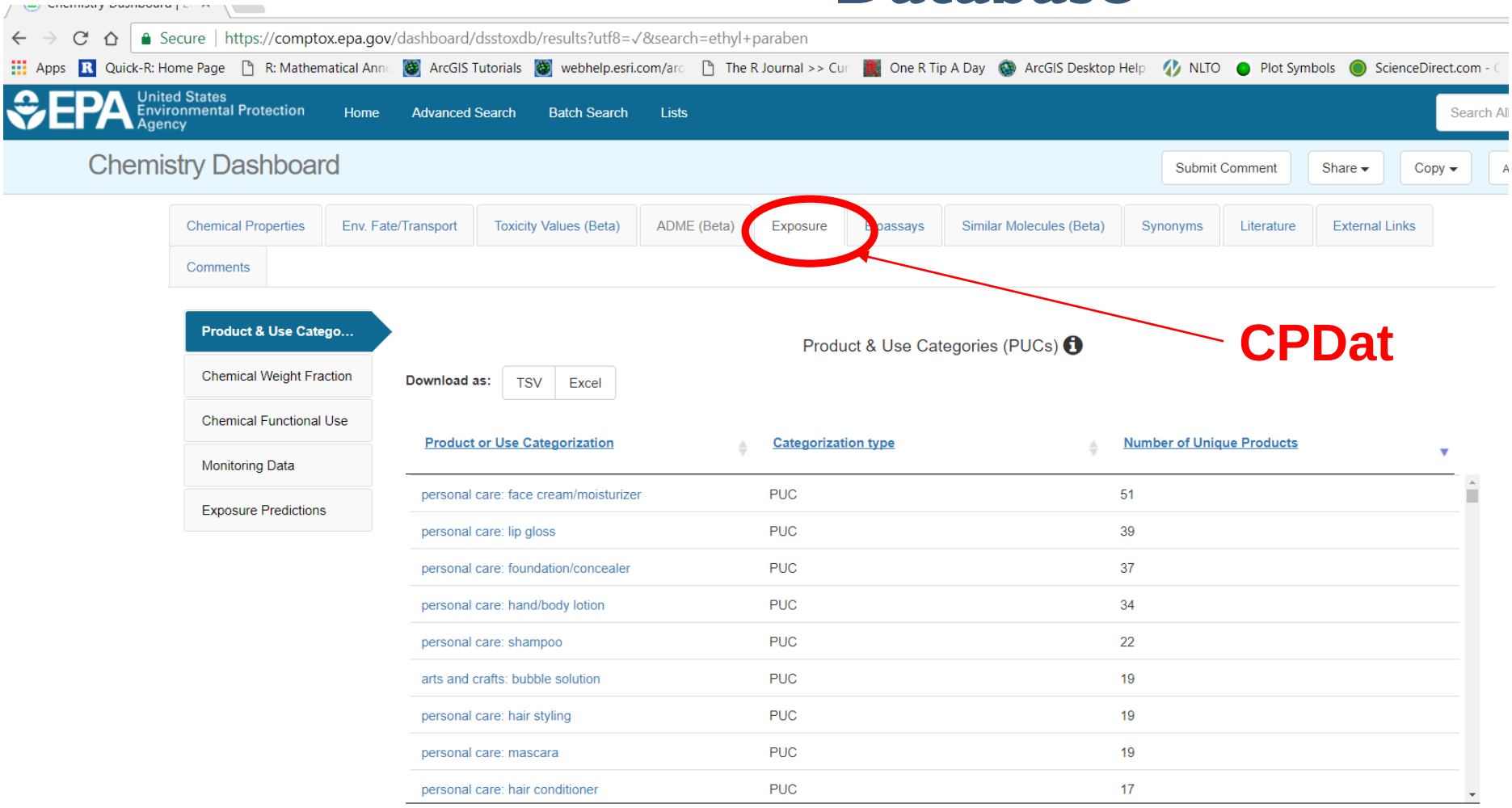
Also available as R Package

Slide from Kristin Isaacs

The Chemistry Dashboard
http://comptox.epa.gov/



Chemicals and Products Database



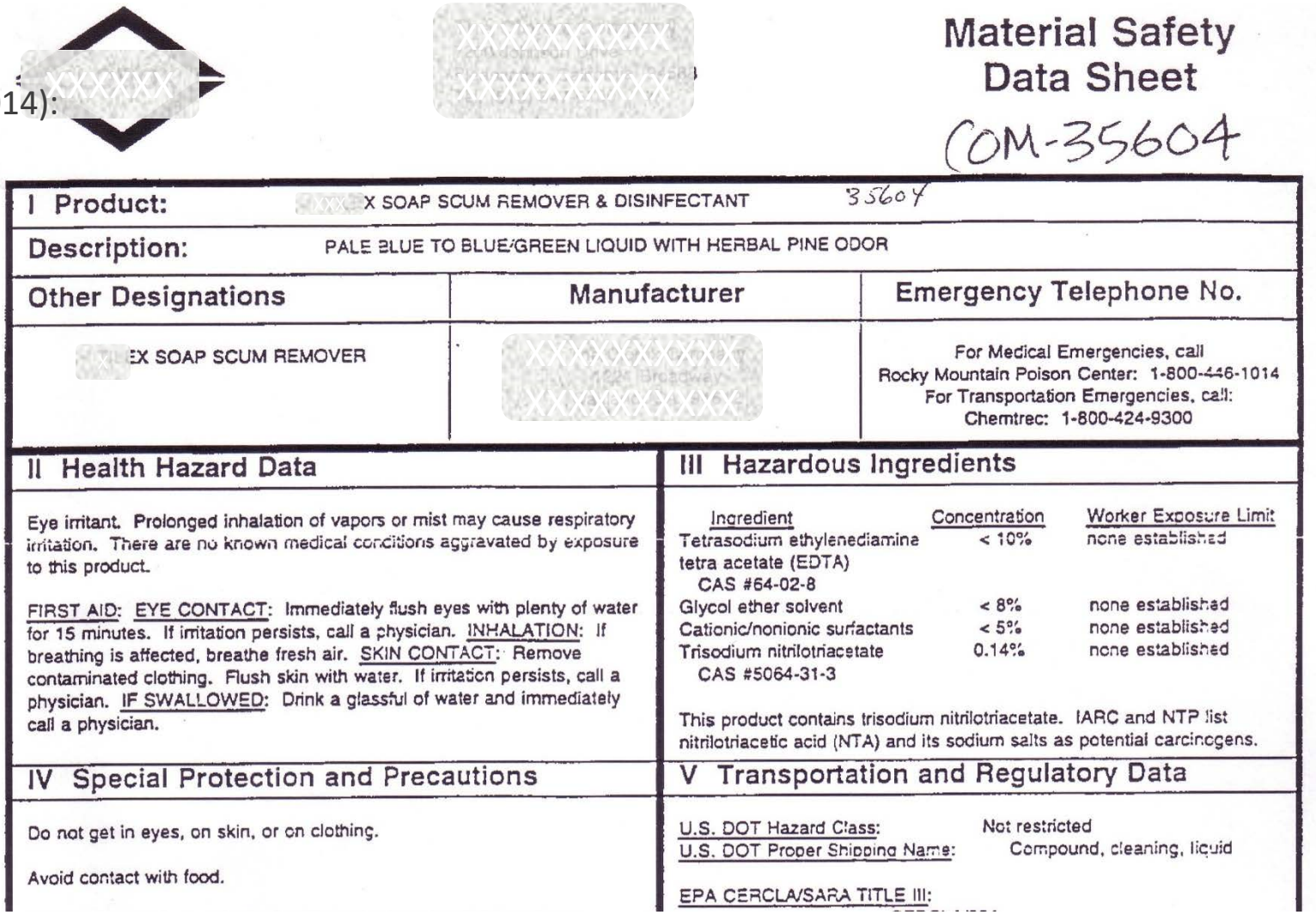
CPCPdb: Material Safety Data Sheets

United States Environmental Protection Agency

Goldsmith et al. (2014):

- ~20,000 product-specific Material Safety Data Sheets (MSDS) curated
- ~2,400 chemicals

Product-specific uses determined using web spider to click through categories (e.g., home goods, bath soaps, baby) to find each product



Material Safety Data Sheet

COM-35604

I Product:	X SOAP SCUM REMOVER & DISINFECTANT 35604	
Description:	PALE BLUE TO BLUE/GREEN LIQUID WITH HERBAL PINE ODOR	
Other Designations	Manufacturer	Emergency Telephone No.
EX SOAP SCUM REMOVER		For Medical Emergencies, call Rocky Mountain Poison Center: 1-800-446-1014 For Transportation Emergencies, call: Chemtrec: 1-800-424-9300

II Health Hazard Data

Eye irritant. Prolonged inhalation of vapors or mist may cause respiratory irritation. There are no known medical conditions aggravated by exposure to this product.

FIRST AID: EYE CONTACT: Immediately flush eyes with plenty of water for 15 minutes. If irritation persists, call a physician. INHALATION: If breathing is affected, breathe fresh air. SKIN CONTACT: Remove contaminated clothing. Flush skin with water. If irritation persists, call a physician. IF SWALLOWED: Drink a glassful of water and immediately call a physician.

III Hazardous Ingredients

Ingredient	Concentration	Worker Exposure Limit
Tetrasodium ethylenediamine tetra acetate (EDTA) CAS #64-02-8	< 10%	none established
Glycol ether solvent	< 8%	none established
Cationic/nonionic surfactants	< 5%	none established
Trisodium nitrilotriacetate CAS #5064-31-3	0.14%	none established

This product contains trisodium nitrilotriacetate. IARC and NTP list nitrilotriacetic acid (NTA) and its sodium salts as potential carcinogens.

IV Special Protection and Precautions

Do not get in eyes, on skin, or on clothing.

Avoid contact with food.

V Transportation and Regulatory Data

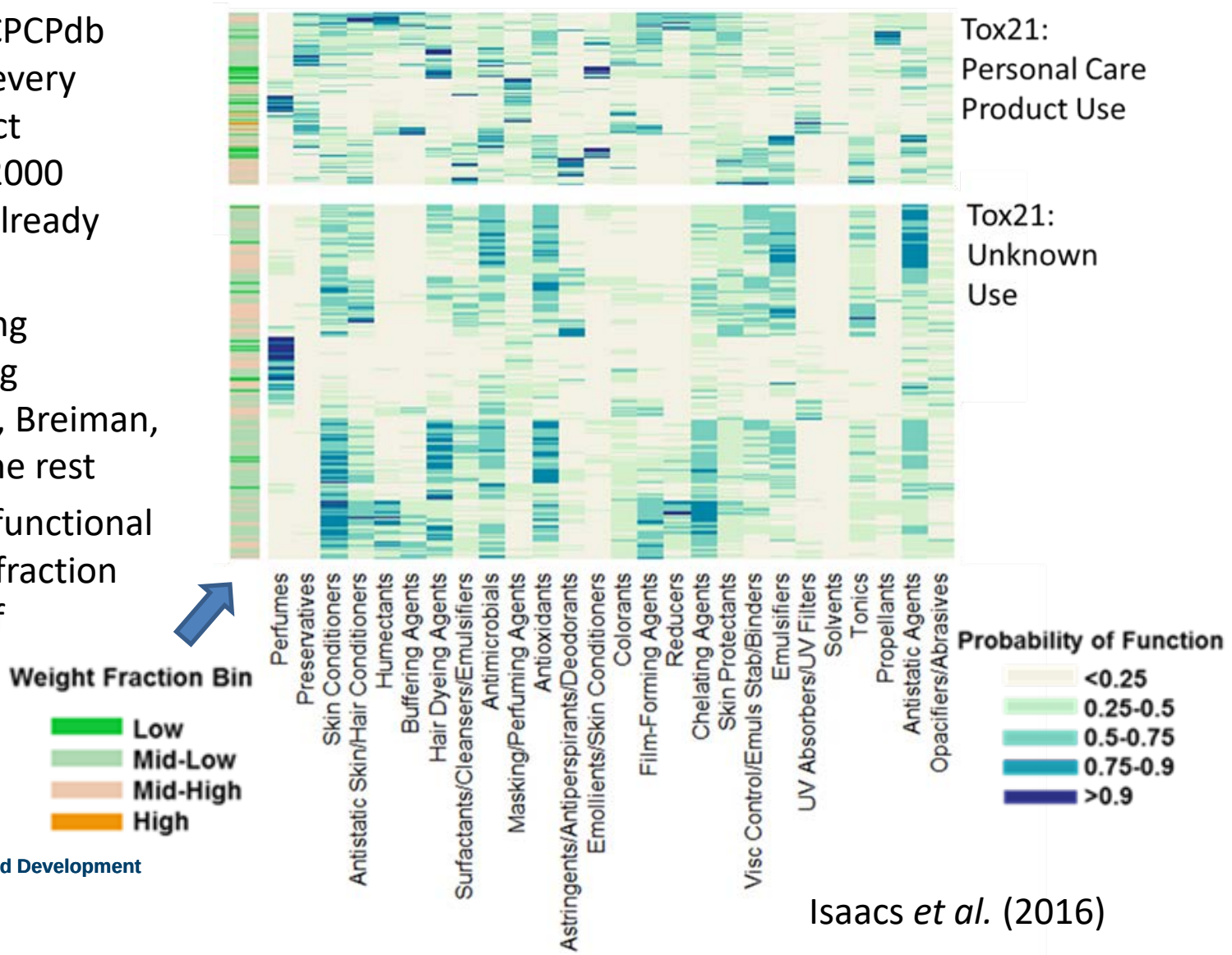
U.S. DOT Hazard Class: Not restricted

U.S. DOT Proper Shipping Name: Compound, cleaning, liquid

EPA CERCLA/SARA TITLE III:

Predicting Chemical Constituents

- Unfortunately, CPCPdb does not cover every chemical-product combination (~2000 chemicals, but already >8000 in Tox21)
- We are now using machine learning (Random Forest, Breiman, 2001) to fill in the rest
- We can predict functional use and weight fraction for thousands of chemicals



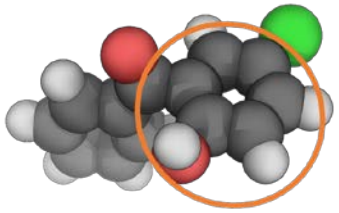
Isaacs *et al.* (2016)

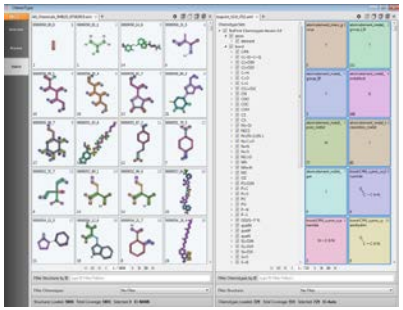
Predicting Function Based on Structure

Random Forest Based Classification Models (Breiman, 2001)

Chemical Structure and Property Descriptors

Use Database (FUSE)



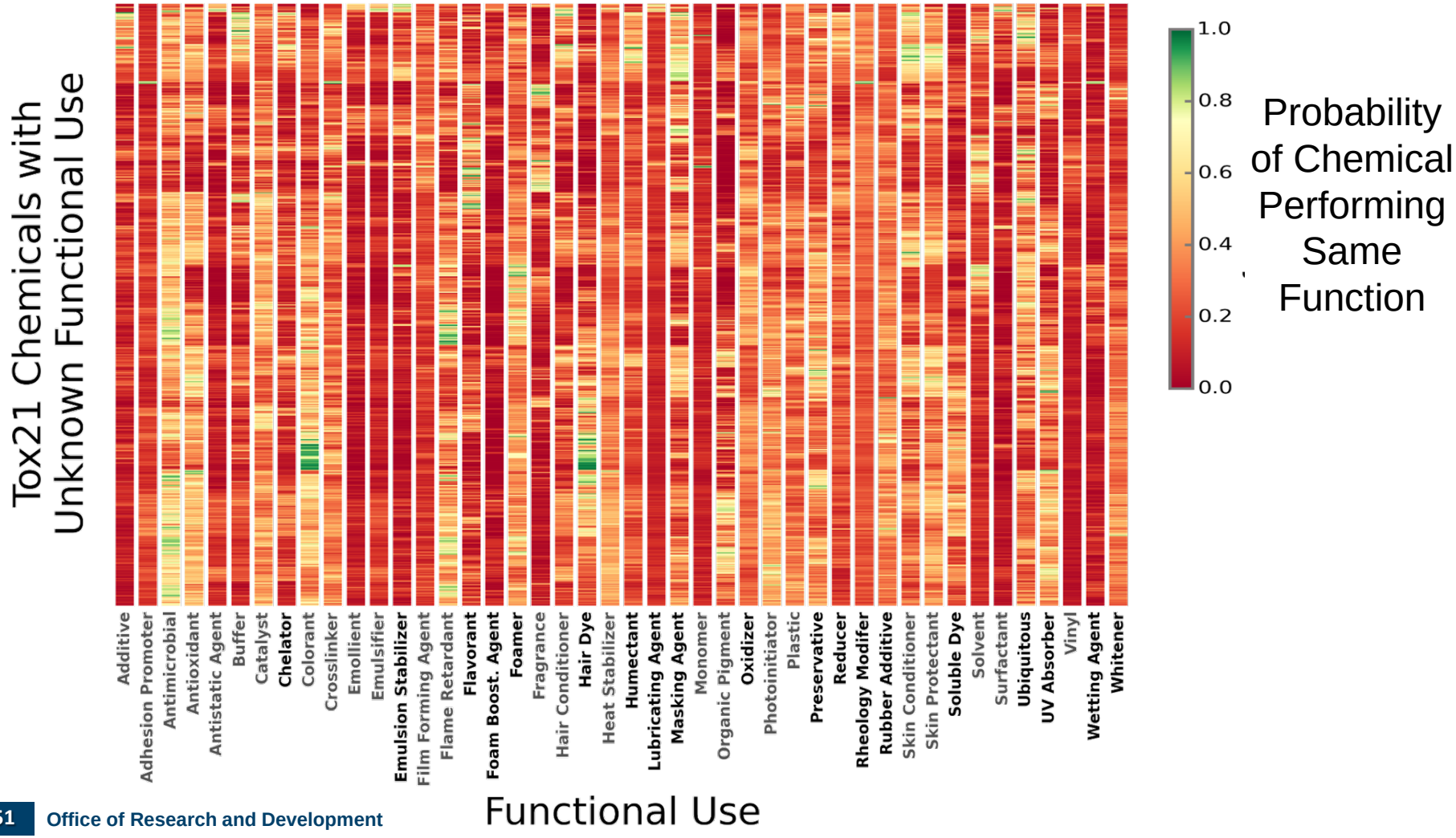


Prediction of Of Potential Alternatives from Chemical Libraries

- Each functional model evaluated on the basis of balanced accuracy, 5-fold CV, and Y-randomization classification errors
- For example, viscosity controllers can be used to thicken or thin out mixtures of chemicals.

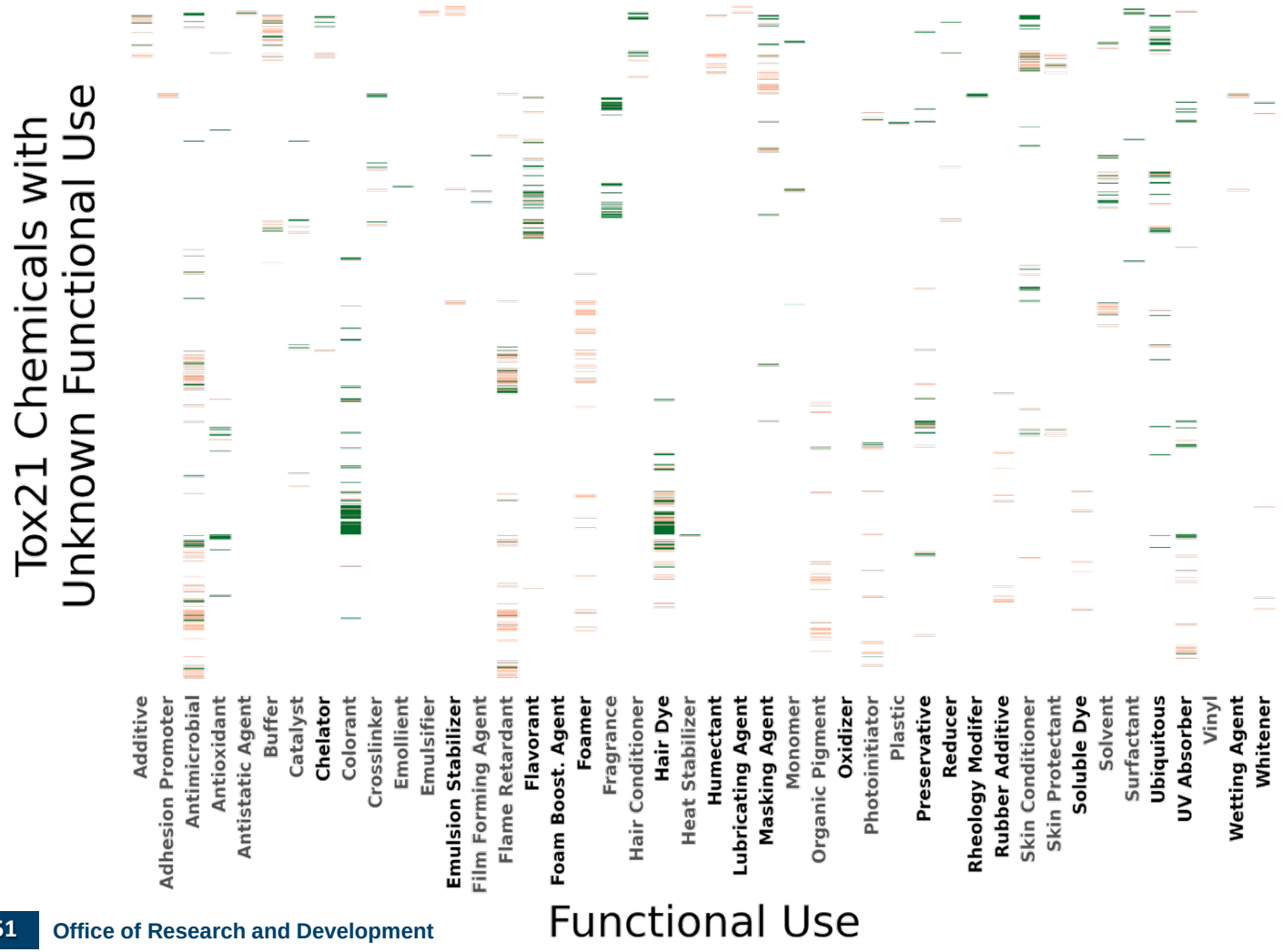
Phillips et al. (2017)

Screening for Alternatives By Function and Bioactivity



Phillips et al. (2017)

Screening for Alternatives By Function and Bioactivity



Phillips et al. (2017)

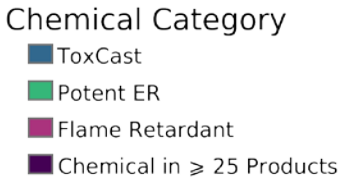
Obtaining New Data with Non-Targeted and Suspect-Screening Analysis

- Not everything is required to have MSDS sheets
- Models present one way forward, but data is always preferable
- New analytic techniques may also allow insight in to the chemical composition of diverse environmental media including household products
- 100 household 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., ES&T just accepted*)

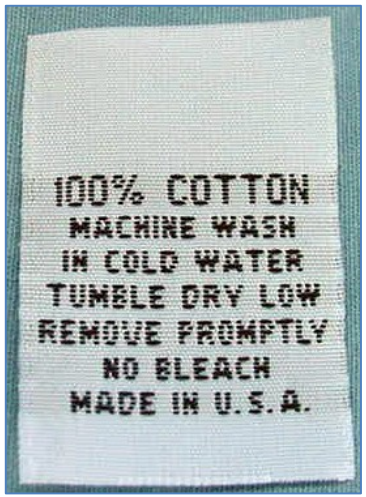
"I'm searching for my keys."

Measuring Chemicals in Household Items

United States Environmental Protection Agency



The chemicals found in a cotton shirt

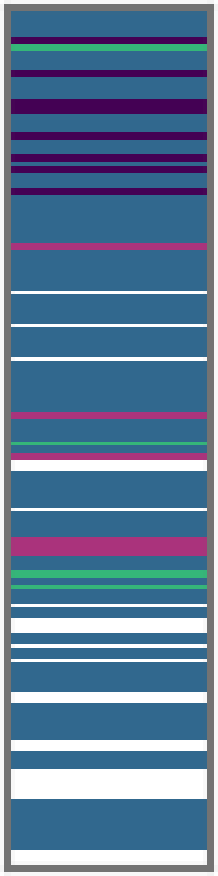


Phillips et al. (*ES&T just accepted*)

Measuring Chemicals in Household Items

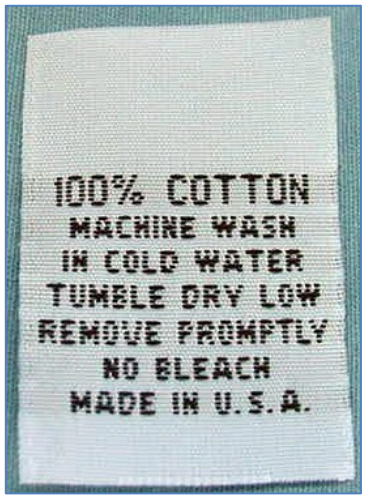
United States Environmental Protection Agency

126 Confirmed



⬅ Chemicals that are present

Chemical Category
- ToxCast
- Potent ER
- Flame Retardant
- Chemical in ⩾ 25 Products

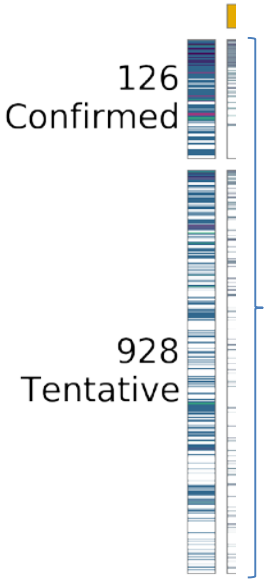


⬅ Chemicals that are absent (but found in other products)

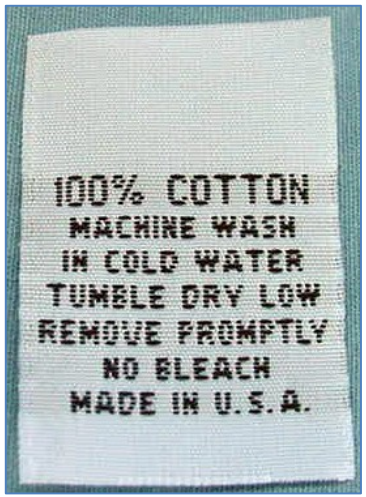
Phillips et al. (*ES&T just accepted*)

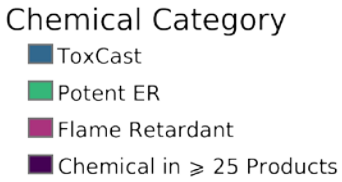
Measuring Chemicals in Household Items

United States Environmental Protection Agency



The chemicals found in a cotton shirt

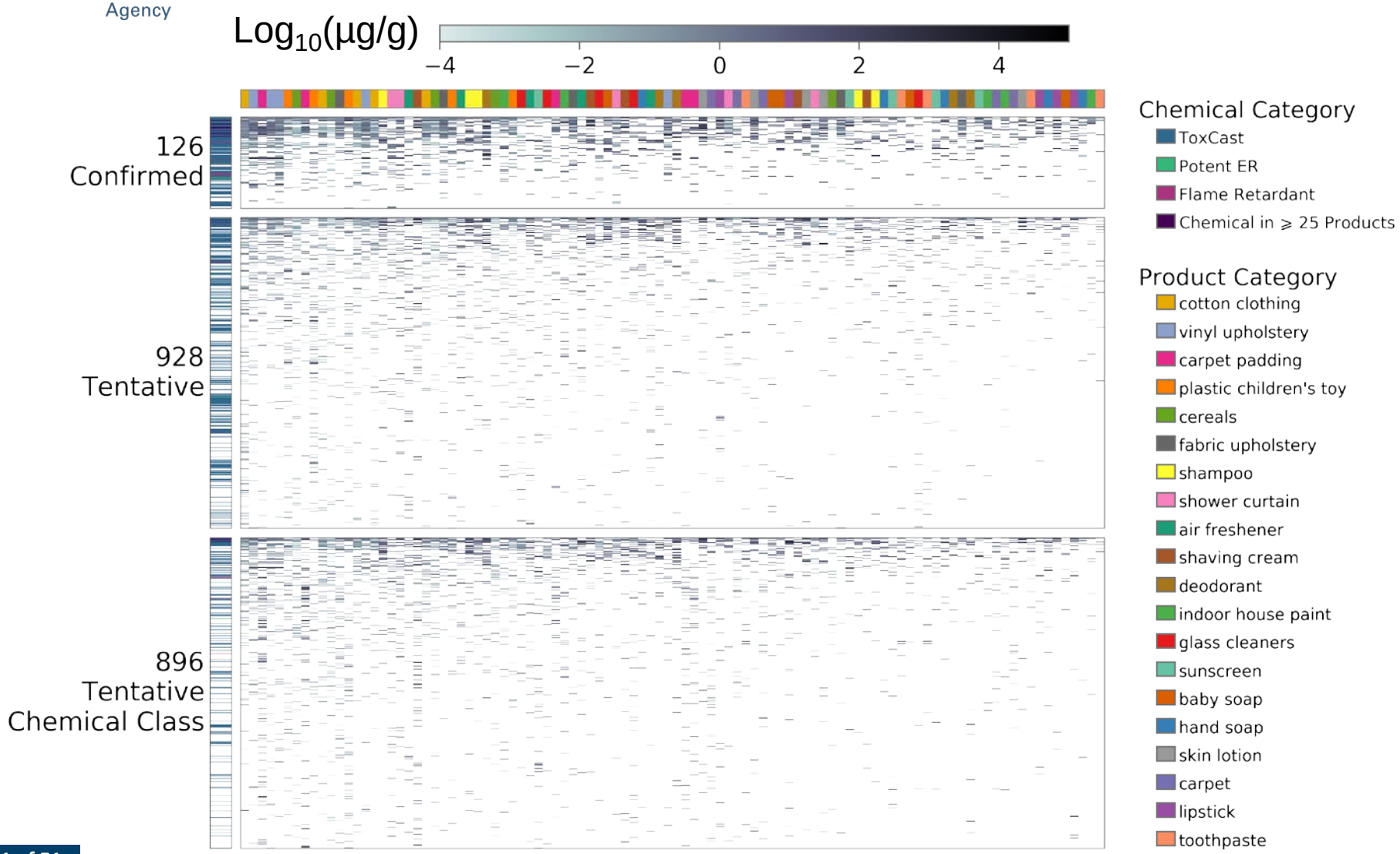




Phillips et al. (*ES&T just accepted*)

Measuring Chemicals in Household Items

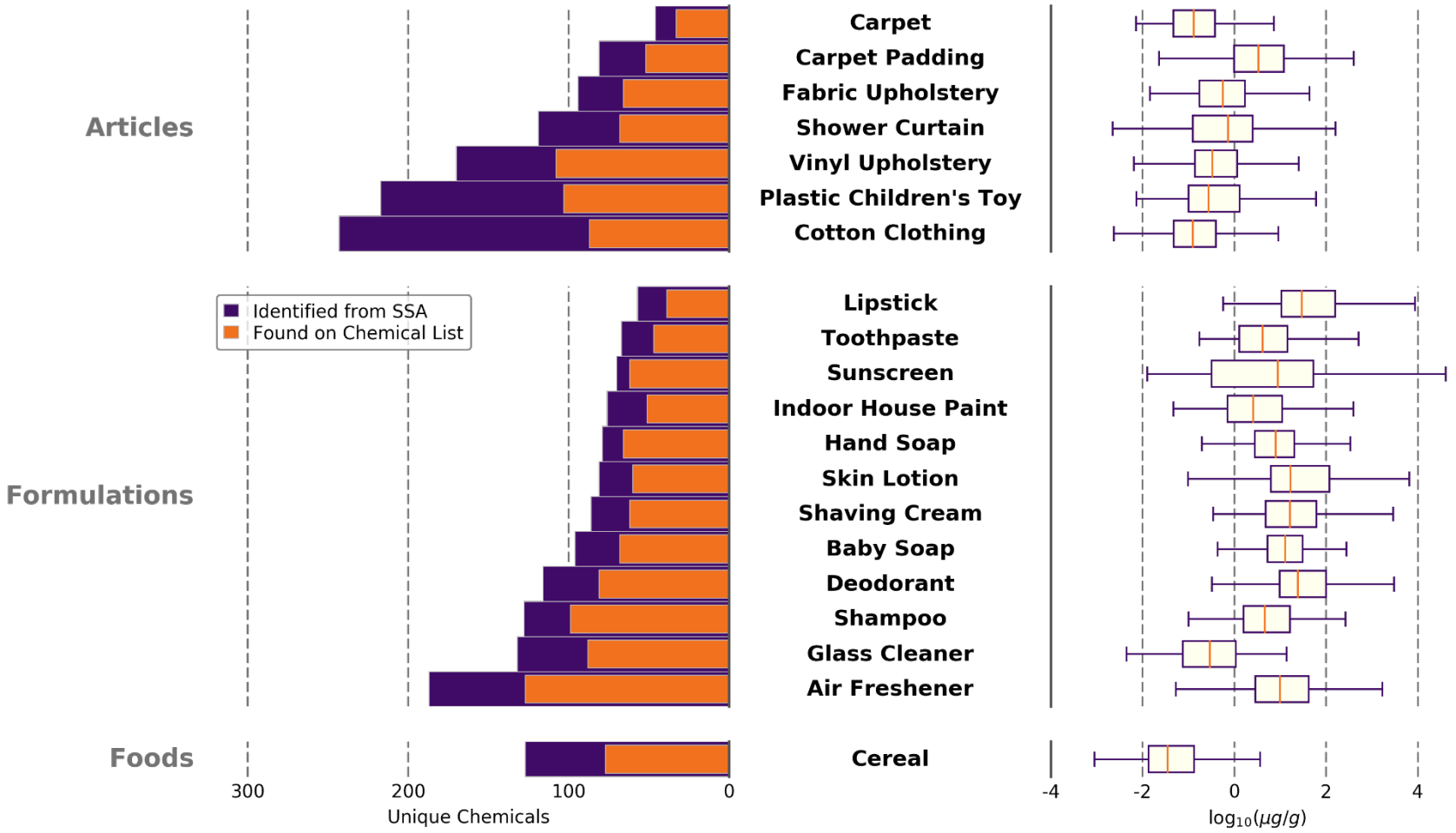
United States Environmental Protection Agency



Phillips et al. (*ES&T just accepted*)

Product Scan Summary

Of 1,632 chemicals confirmed or tentatively identified, 1,445 were not present in CPCPdb (Goldsmith, et al., 2015)



Phillips et al. (*ES&T just accepted*)

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

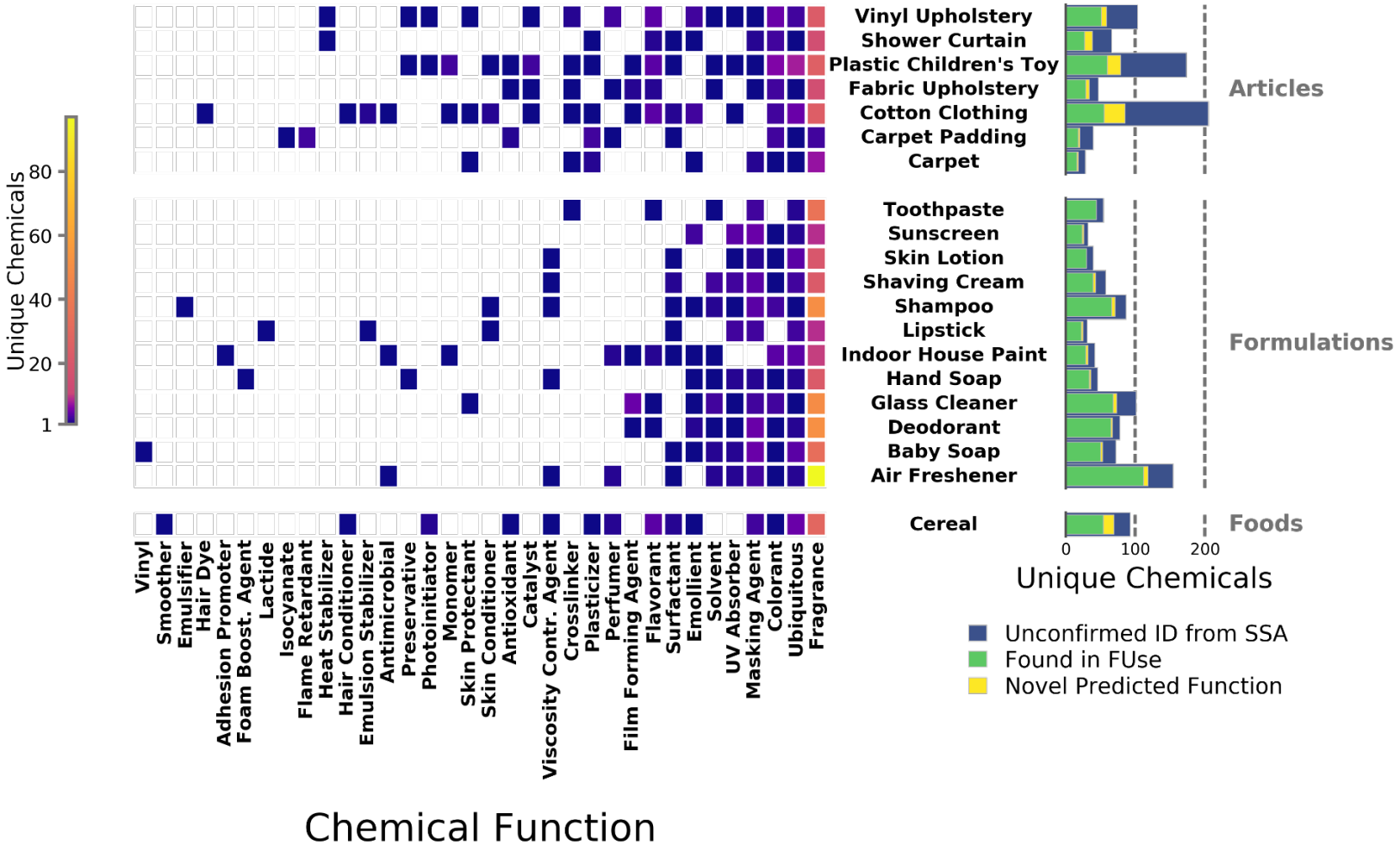
"As chemists we are obliged to accept the assignment of barium to the observed activity, but as nuclear chemists working very closely to the field of physics we cannot yet bring ourselves to take such a drastic step, which goes against all previous experience in nuclear physics. It could be, however,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

Predicting Chemical Function

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



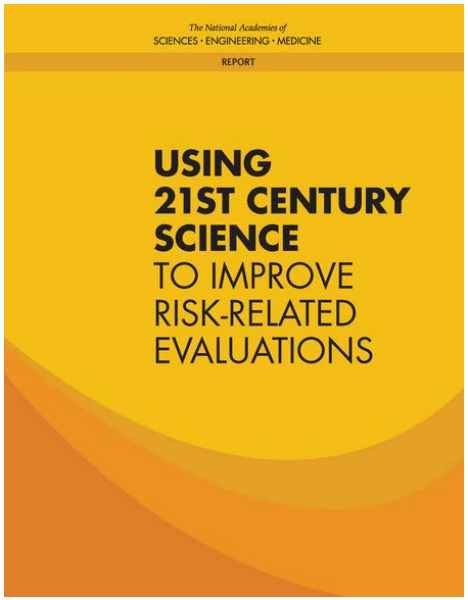
Phillips et al. (*ES&T just accepted*)

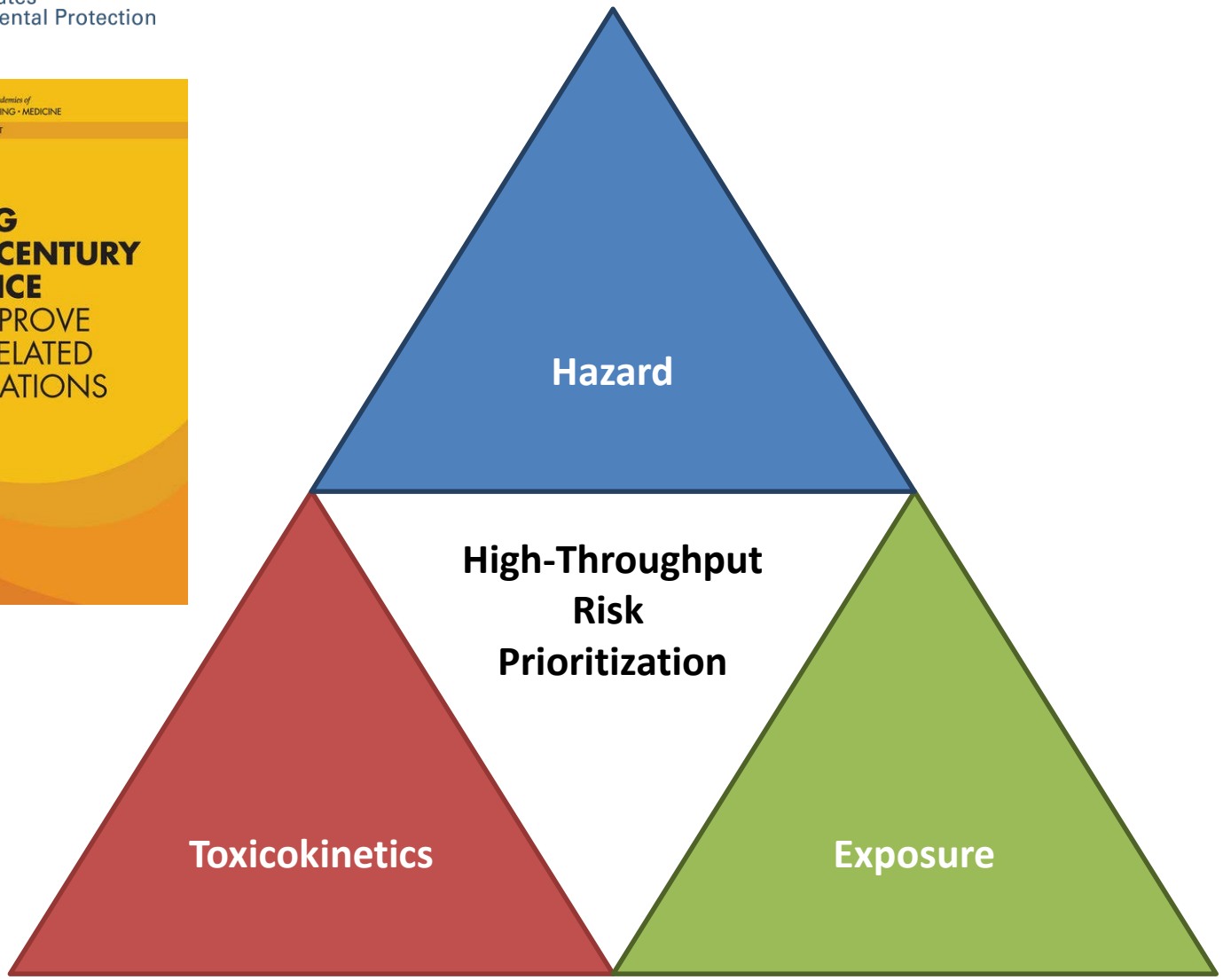
Caveats to Non-Targeted Screening

- **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**

Exposure-Based Priority Setting

United States Environmental Protection Agency





The Structure of Chemical Exposure

United States Environmental Protection Agency

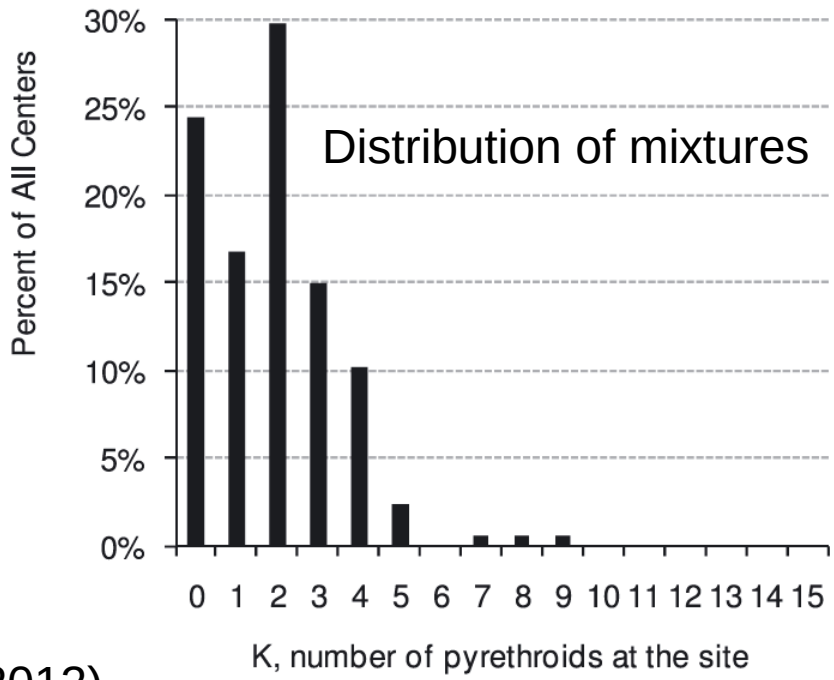
finch species

	Cuba	Hispanolia	Jamaica	Puerto Rico	Guadeloupe	
Loxigilla noxis	0	0	0	0	1	**1**
Melanospiza richardsoni	0	0	0	0	0	**0**
Tiara olivacea	1	1	1	1	0	**4**
Tiara bicolor	0	1	1	1	1	**4**
Tiara canora	1	0	0	0	0	**1**
Loxipasser anoxanthus	0	0	1	0	0	**1**
	2	**2**	**3**	**2**	**2**	**11**

chemical species

	site 1	site 2	site 3	site 4	site 5	
chemical 1	1	0	0	0	1	**2**
chemical 2	1	1	0	1	1	**4**
chemical 3	0	0	0	0	1	**1**
chemical 4	1	0	0	0	0	**1**
chemical 5	0	0	1	0	0	**1**
chemical 6	0	0	1	0	0	**1**
	3	**1**	**2**	**1**	**3**	**10**

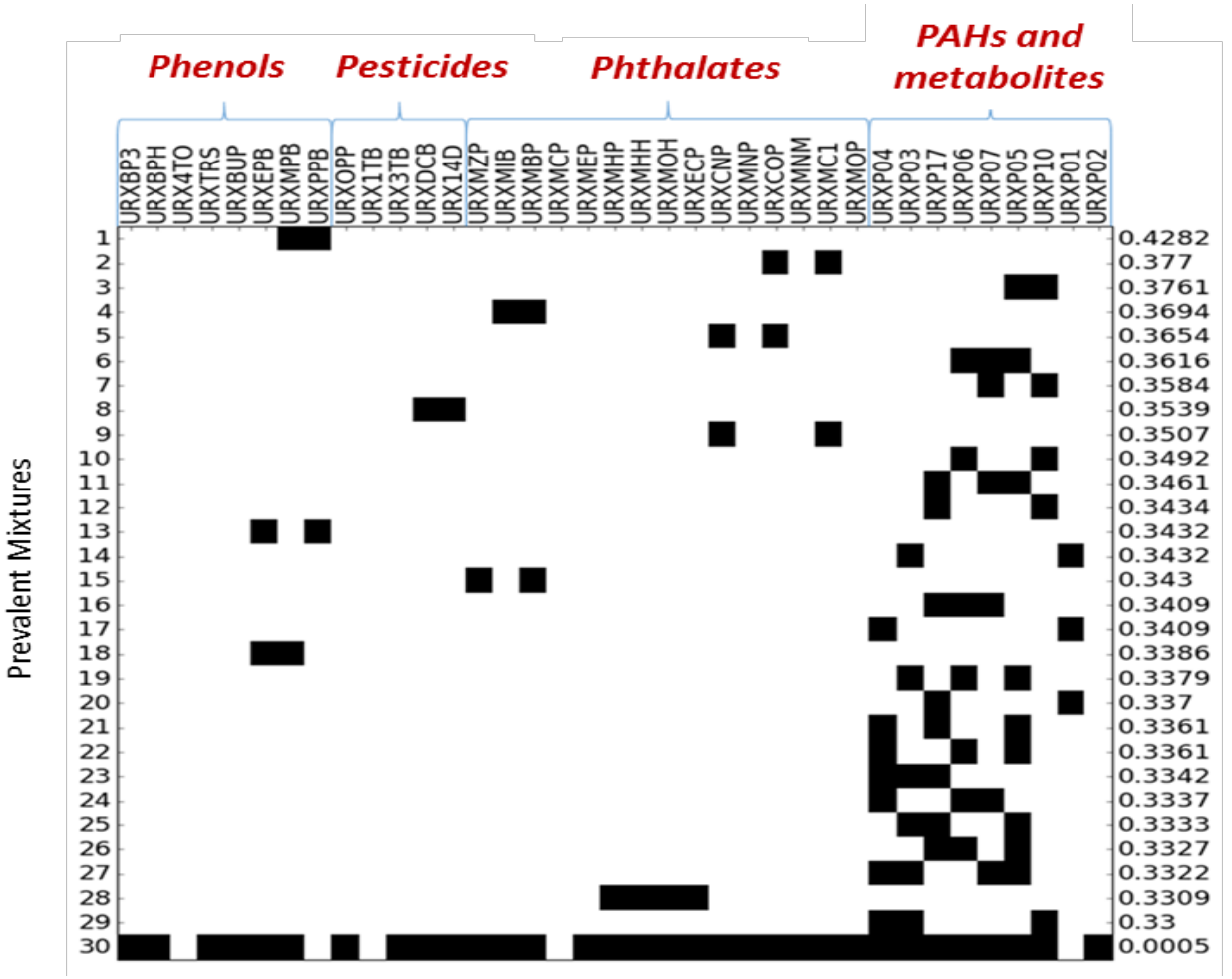
- For **n** chemicals 2^n combinations are possible
  - However, not all are observed
- Diamond (1975): Not all finch species present on all islands of Caribbean
- Tornero-Velez et al. (2012): Not all chemical combinations present at all sites



Tornero-Velez et al. (2012)

Identifying Prevalent Mixtures

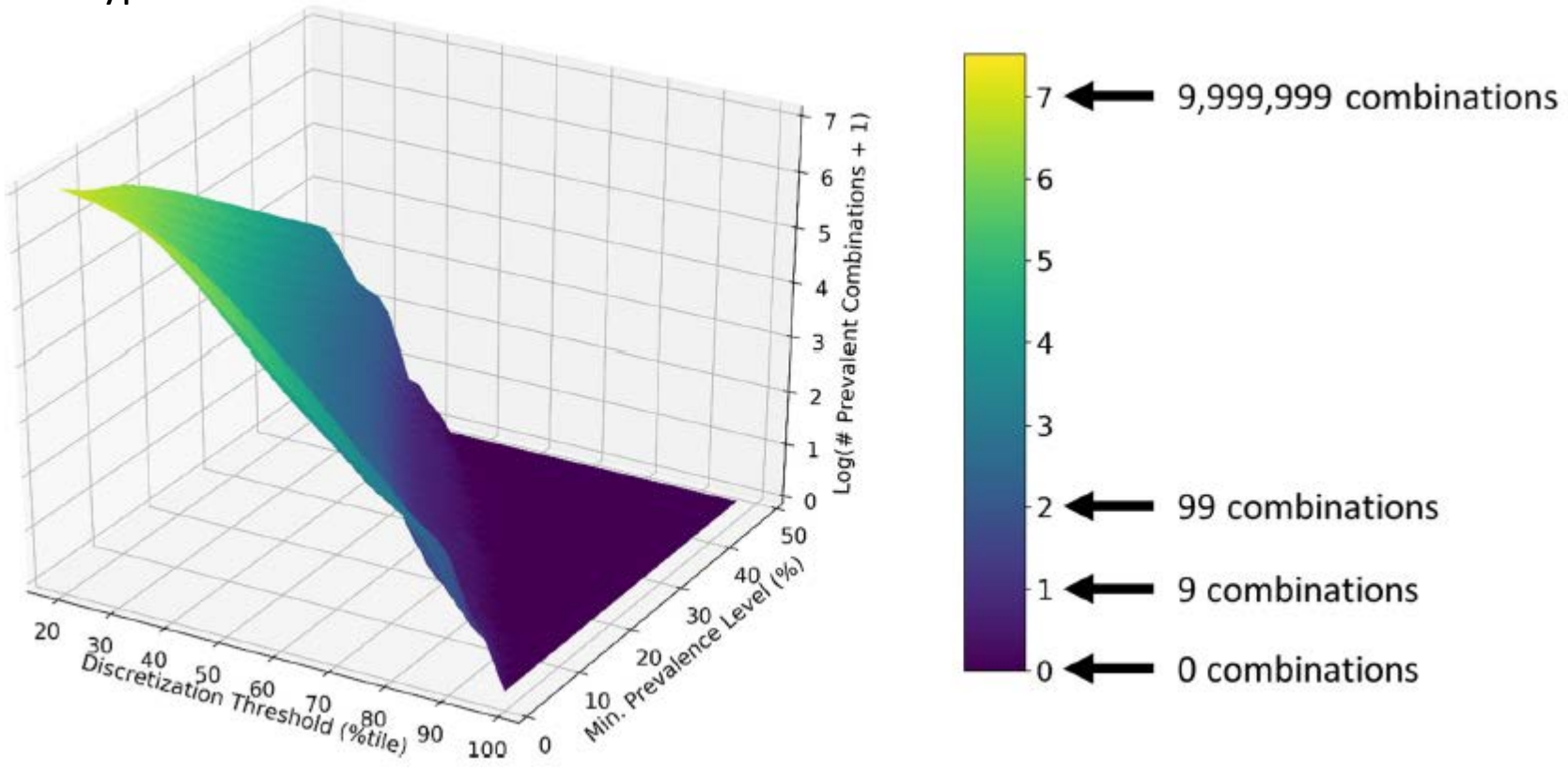
- Kapraun et al. (2017) used frequent itemset mining (FIM, Borgelt, 2012) to identify combinations of items (chemicals) that co-occur together within CDC NHANES samples from same individual
- Used total population median concentration as threshold for “presence”
- Identified a few dozen mixtures present in >30% of U.S. population



Kapraun et al. (2017)

A Testable Number of Combinations

While high throughput screening (HTS) allows thousands of tests, there are millions of hypothetical combinations



"Exposure based priority setting" (NAS, 2017) allows identification of most important mixtures to test

Kapraun et al. (2017)

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 one path forward for identifying potential hazard, but the real world is complicated by toxicokinetics, mixtures, variability (and more)

- Using *in vitro* methods developed for pharmaceuticals, we can make useful predictions of TK for large numbers of chemicals

- Exposure data key to risk-based prioritization
  - Consensus modeling provides one path forward, but only as good as available data (at best)
  - New analytical chemistry tools (i.e., non-targeted analysis or NTA) may provide the data needed to understand what and how we are exposed to

- Exposure-based priority setting allows identification of the most relevant mixtures

EPA United States Environmental Protection Agency

Chemical Safety for Sustainability (CSS) Rapid Exposure and Dosimetry (RED) Project

NCCT
Chris Grulke
Greg Honda*
Richard Judson
Andrew McEachran*
Robert Pearce*
Ann Richard
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

NERL
Craig Barber
Namdi Brandon*
Peter Egeghy
Hongtai Huang*
Brandall Ingle*
Kristin Isaacs
Seth Newton
Katherine Phillips
Paul Price
Jeanette Reyes*
Jon Sobus
John Streicher*
Mark Strynar
Mike Tornero-Velez
Elin Ulrich
Dan Vallero
Barbara Wetmore

*Trainees

Lead CSS Matrix Interfaces:
John Kenneke (NERL)
John Cowden (NCCT)

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
Research Triangle Institute
Timothy Fennell
ScitoVation
Harvey Clewell
Kamle Mansouri
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
Hyeong-Moo Shin
University of Michigan
Olivier Jolliet
University of North Carolina, Chapel Hill
Alex Tropsha

United States Environmental Protection Agency

References

- Borgelt, C. (2012). Frequent item set mining. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2(6), 437-456.
- Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.
- Breyer, Stephen. Breaking the vicious circle: Toward effective risk regulation. Harvard University Press, 2009
- Diamond JM. Assembly of species communities. Pp. 342–444 in Cody ML, Diamond JM (eds). Ecology and Evolution of Communities. Cambridge, MA: Belkap Press, 1975.
- Egeghy, P. P., et al. (2012). The exposure data landscape for manufactured chemicals. Science of the Total Environment, 414, 159-166.
- 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. (2014). Development of a consumer product ingredient database for chemical exposure screening and prioritization. Food and chemical toxicology, 65, 269-279.
- Hahn, O., & Straßmann, F. (1938). Über die Entstehung von Radiumisotopen aus Uran durch Bestrahlen mit schnellen und verlangsamten Neutronen. Naturwissenschaften, 26(46), 755-756.
- Kaewkhaw, R., et al. (2016). Treatment paradigms for retinal and macular diseases using 3-D retina cultures derived from human reporter pluripotent stem cell linestreatment design using PSC-Derived 3-D retina cultures. Investigative ophthalmology & visual science, 57(5), ORSFI1-ORSFI11.
- Kapraun, Dustin et al., “A Method for Identifying Prevalent Chemical Combinations in the US Population,” Environmental Health Perspectives, 2017
- Kavlock, Robert, et al. "Update on EPA’s ToxCast program: providing high throughput decision support tools for chemical risk management." Chemical research in toxicology 25.7 (2012): 1287-1302.
- MacLeod, Matthew, et al. "The state of multimedia mass-balance modeling in environmental science and decision-making." (2010): 8360-8364
- National Academies of Sciences, Engineering, and Medicine. (2017). Using 21st century science to improve risk-related evaluations. National Academies Press.
- National Research Council. (1983). Risk Assessment in the Federal Government: Managing the Process Working Papers. National Academies Press.
- National Research Council. (2007) Toxicity testing in the 21st century: a vision and a strategy. National Academies Press.
- O’Connell, S. G., Kincl, L. D., & Anderson, K. A. (2014). Silicone wristbands as personal passive samplers. Environmental science & technology, 48(6), 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, 2017
- 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”, accepted at Environmental Science & Technology.
- Rotroff, Daniel, et al., (2010) “Incorporating human dosimetry and exposure into high-throughput in vitro toxicity screening.” Tox. Sciences 117(2), 348-58
- Schmidt, Charles W. "TOX 21: new dimensions of toxicity testing." Environmental health perspectives 117.8 (2009): A348.
- Shin, Hyeong-Moo, et al. "Risk-based high-throughput chemical screening and prioritization using exposure models and in vitro bioactivity assays." Environmental science & technology 49.11 (2015): 6760-6771.
- 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)
- Tornero-Velez et al. (2012) “Biogeographical Analysis of Chemical Co-Occurrence to Identify Priorities for Mixtures Research”: Risk Analysis, 32(2) 224-236
- 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. “Evaluating In Vitro-In Vivo Extrapolation" accepted at Toxicological Sciences,.
- Wang, Y.-H. (2010). “Confidence Assessment of the Simcyp Time-Based Approach and a Static Mathematical Model in Predicting Clinical Drug-Drug Interactions for Mechanism-Based CYP3A Inhibitors.” Drug Metabolism and Disposition 38(7), 1094-1104
- 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.