High Throughput Exposure Science for Chemical Decision Making

John F. Wambaugh

National Center for Computational Toxicology
Office of Research and Development
United States Environmental Protection Agency
Research Triangle Park, North Carolina 27711

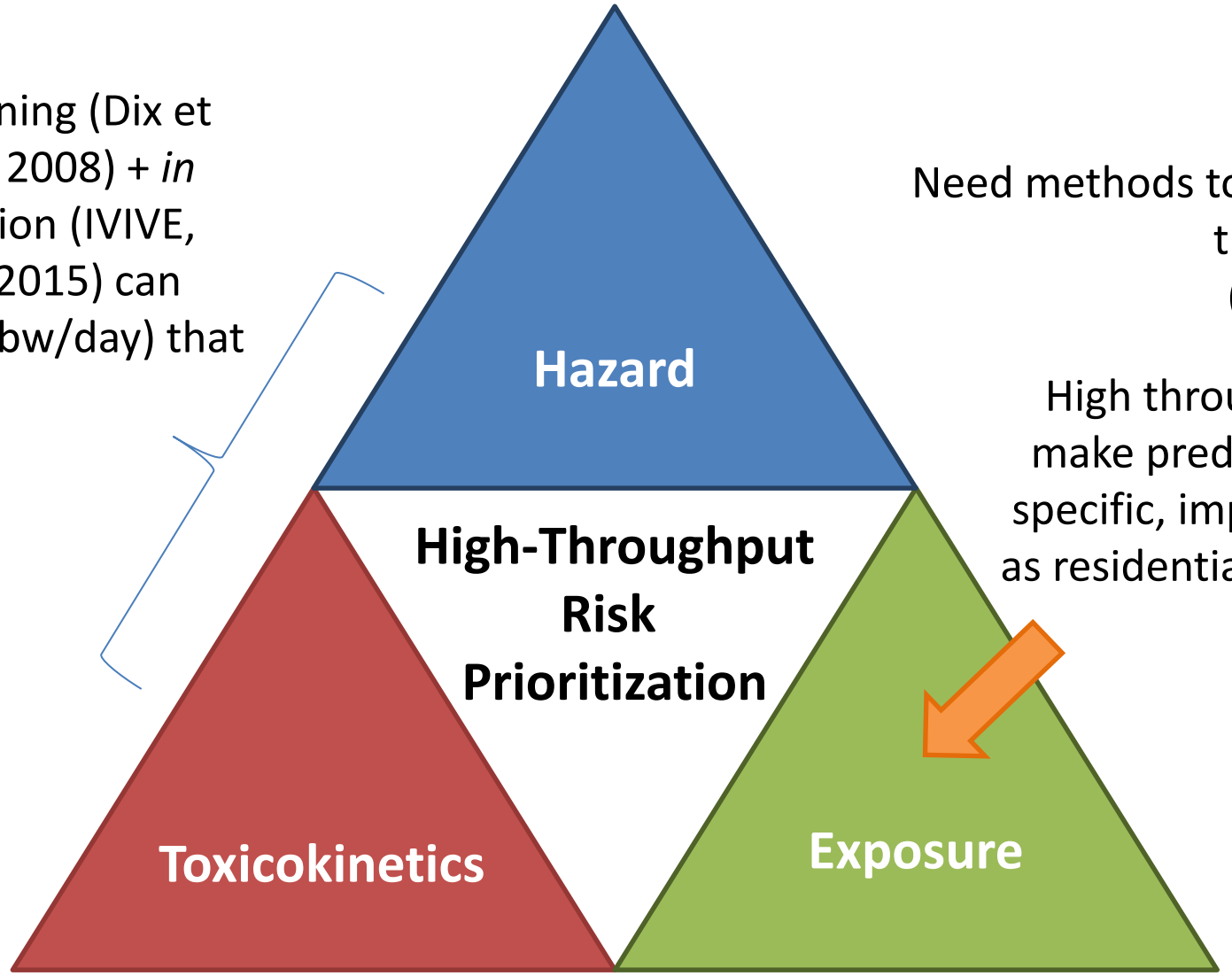
The views expressed in this presentation are those of the author and do not necessarily reflect the views or policies of the U.S. EPA

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

Total Exposure Health 2018: "Bridging Human Exposure & Precision Medicine"
September 7, 2018

Risk = Hazard x Exposure

High throughput screening (Dix et al., 2007, Collins et al., 2008) + *in vitro-in vivo* extrapolation (IVIVE, Wetmore et al., 2012, 2015) can predict a dose (mg/kg bw/day) that might be adverse



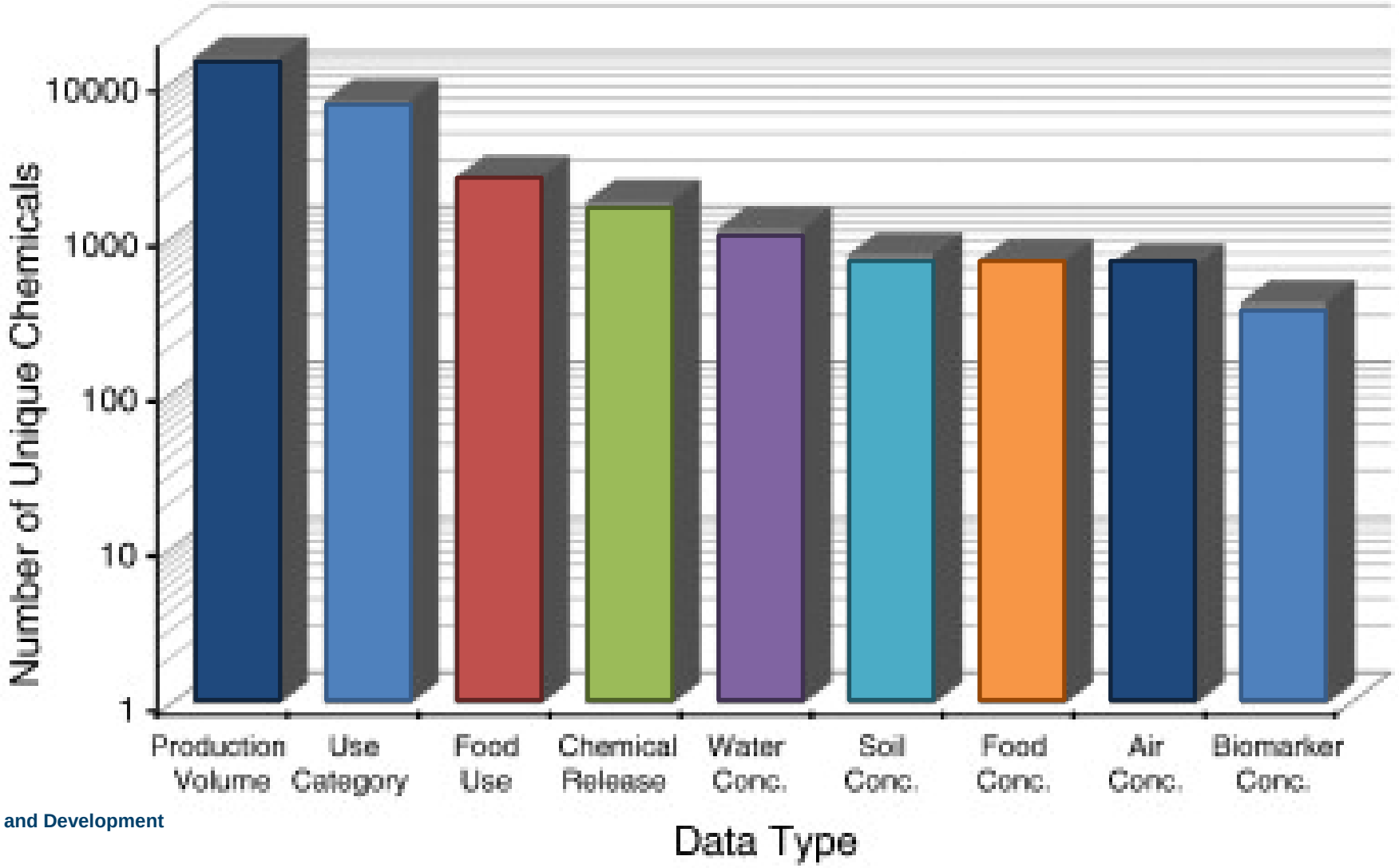
Need methods to forecast exposure for thousands of chemicals (Wetmore et al., 2015)

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

NRC (1983)

Limited Available Data for Exposure Estimation

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



Biomonitoring Data

The National Health and Nutrition Examination Survey (NHANES) provides targeted biomonitoring data of chemicals and metabolites in human blood and urine

Chemicals in the *Fourth Report: Updated Tables, March 2018*

CDC (2018)

CDC's *Fourth National Report on Human Exposure to Environmental Chemicals: Updated Tables* provides exposure data on the following chemicals or classes of chemicals. The *Updated Tables* contain cumulative data from national samples collected beginning in 1999–2000 and as recently as 2015-2016. Not all chemicals were measured in each national sample. The data tables are available at http://www.cdc.gov/exposurereport. An asterisk (*) indicates the chemical has been added since publication of the *Fourth Report* in 2009.

Phthalate and Phthalate Alternative Metabolites

Mono-benzyl phthalate (MBzP)
Mono-3-hydroxybutyl phthalate (MHBP)*
Mono-n-butyl phthalate (MnBP)
Mono-2-methyl-2-hydroxypropyl phthalate (MHiBP)*
Mono-isobutyl phthalate (MiBP)
Mono-cyclohexyl phthalate (MCHP)
Mono-ethyl phthalate (MEP)
Mono-2-ethylhexyl phthalate (MEHP)
Mono-(2-ethyl-5-hydroxyhexyl) phthalate (MEHHP)

Organophosphorus Insecticides: Dialkyl Phosphate Metabolites

Diethylphosphate (DEP)
Dimethylphosphate (DMP)
Diethylthiophosphate (DETP)
Dimethylthiophosphate (DMTP)
Diethyldithiophosphate (DEDTP)
Dimethyldithiophosphate (DMDTP)

Pyrethroid Metabolites

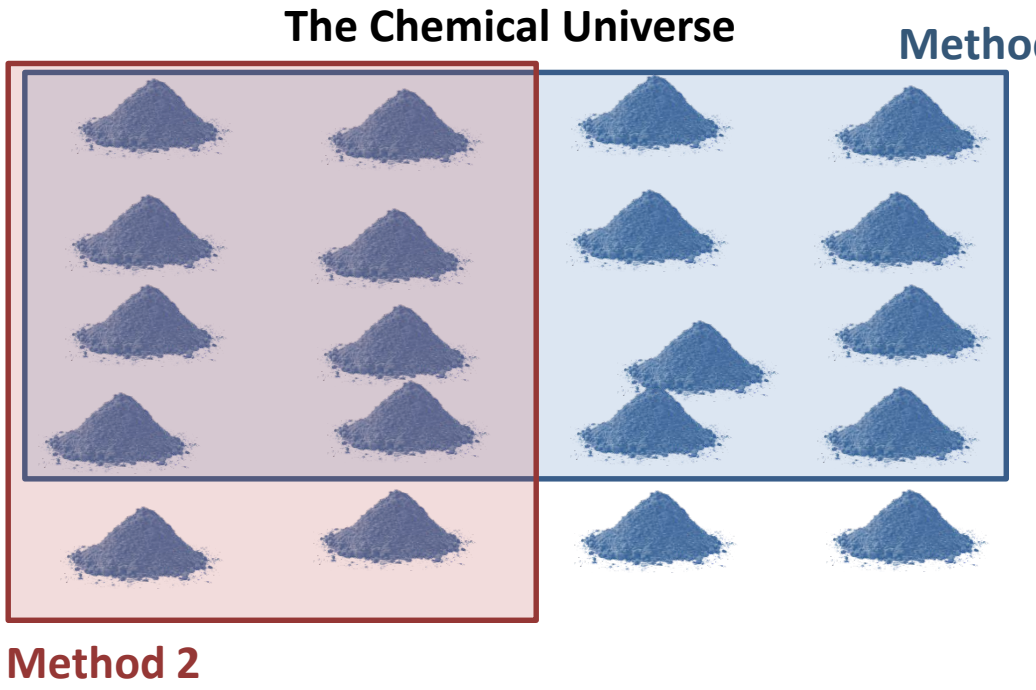
trans-3-(2,2-Dichlorovinyl)-2,2-dimethylcyclopropane carboxylic acid (trans-DCCA)
cis-3-(2,2-Dibromovinyl)-2,2dimethylcyclopropane carboxylic acid (cis-DBCA)
4-Fluoro-3-phenoxy-benzoic acid*
3-Phenoxybenzoic acid*

There are hundreds of chemicals, and yet Park *et al.* (2012) and others have seen evidence for many others

EPA's Non-Targeted Analysis Collaborative Trial (ENTACT)

Suspect screening / Non-targeted analyses (SSA/NTA) present opportunities for new exposure data

What NTA methods are available? What is the coverage of chemical universe and matrices? How do methods differ in their coverage?



Led by Jon Sobus and Elin Ulrich (EPA/NERL)

Phase 1:

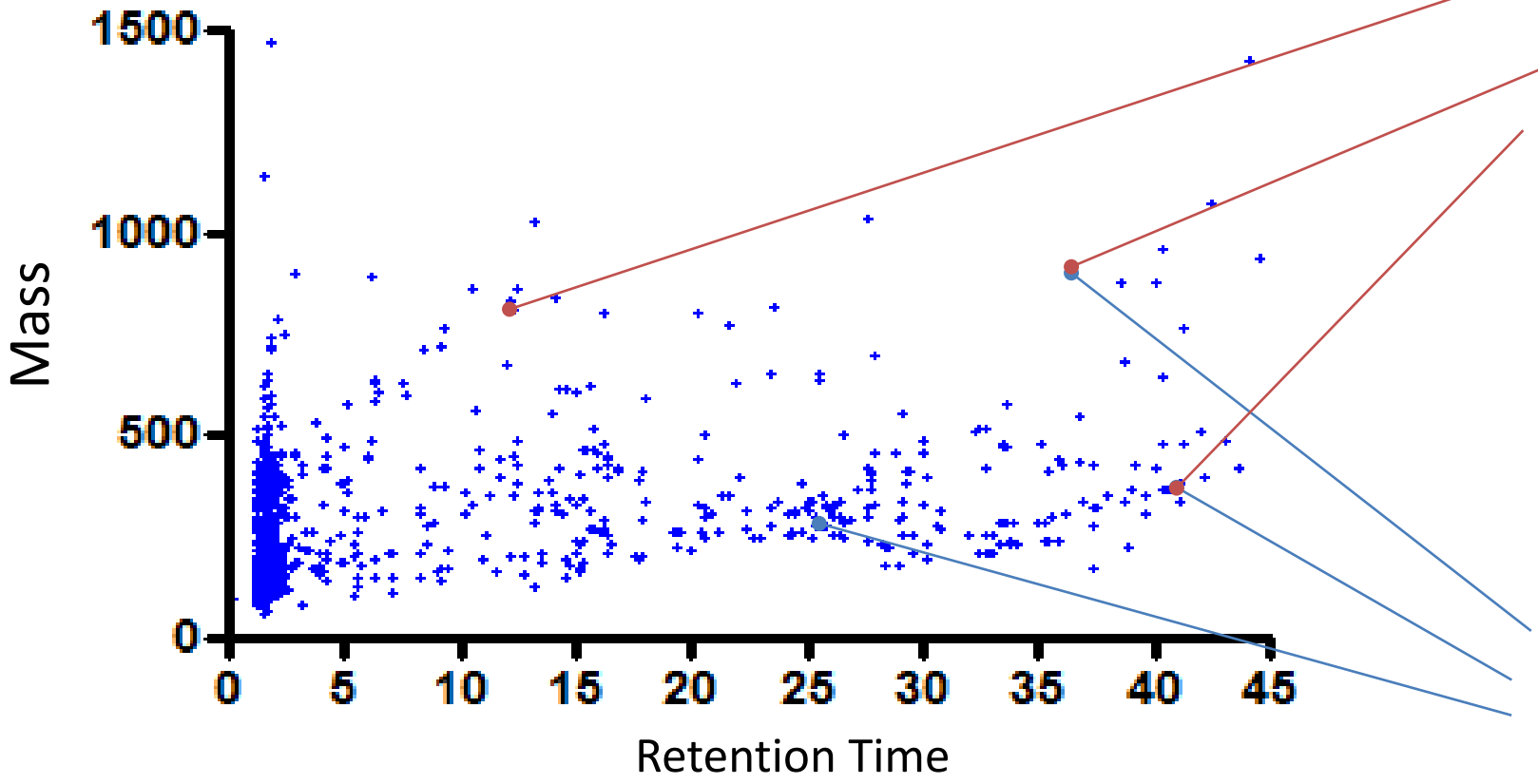
- Collaborators provided 10 mixtures of 100-400 ToxCast chemicals each
- MS vendors provided with individual chemical standards

Phase 2: Fortified reference house dust, human serum, and silicone wristbands

Sobus et al. (2017)

What We Learned from Suspect Screening in House Dust

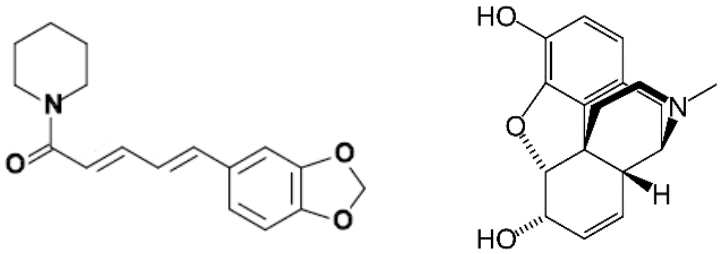
947 Peaks in an American Health Homes Dust Sample



Liquid chromatography peaks 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:

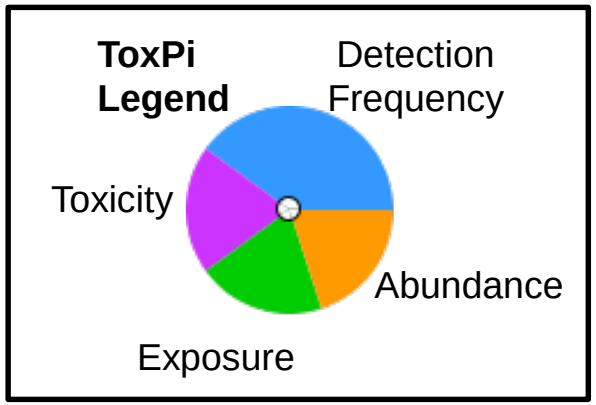


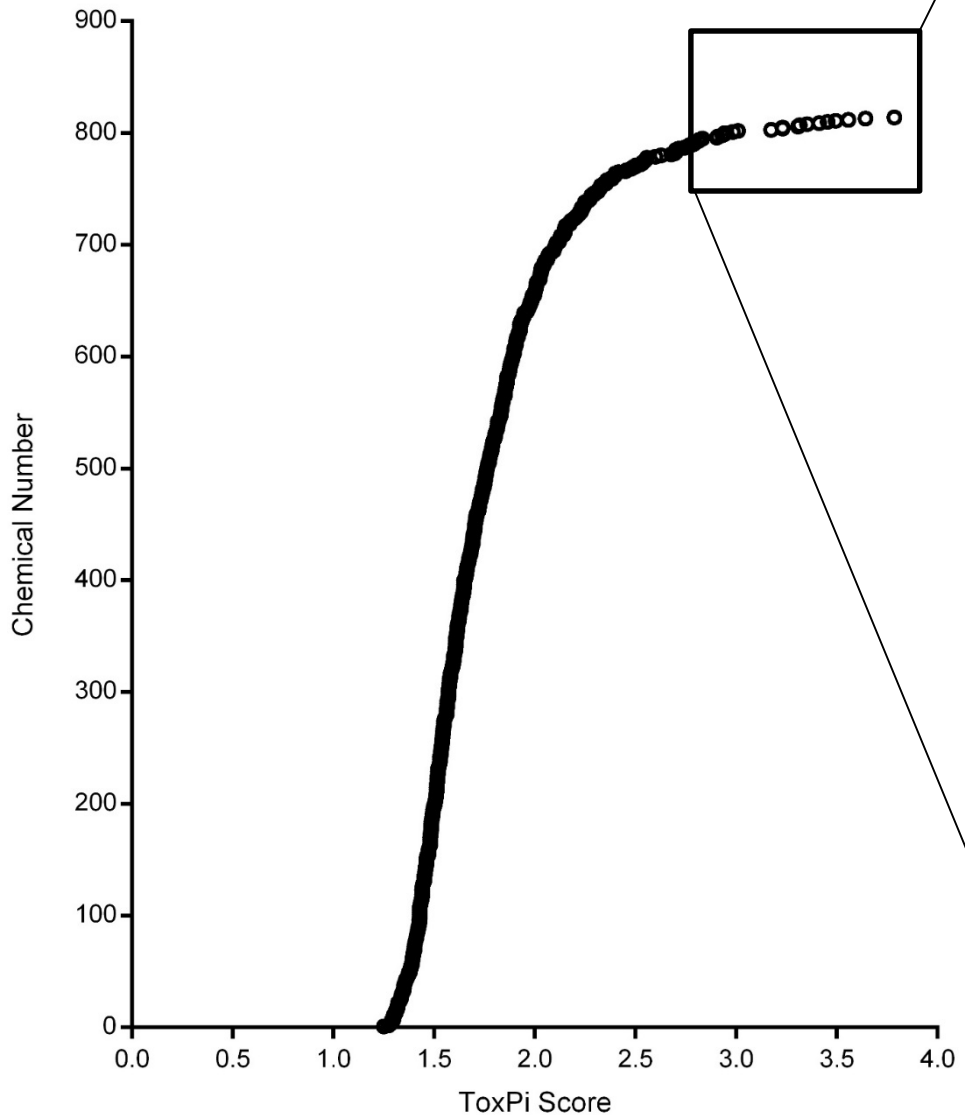
Is chemical A present, chemical B, or both?

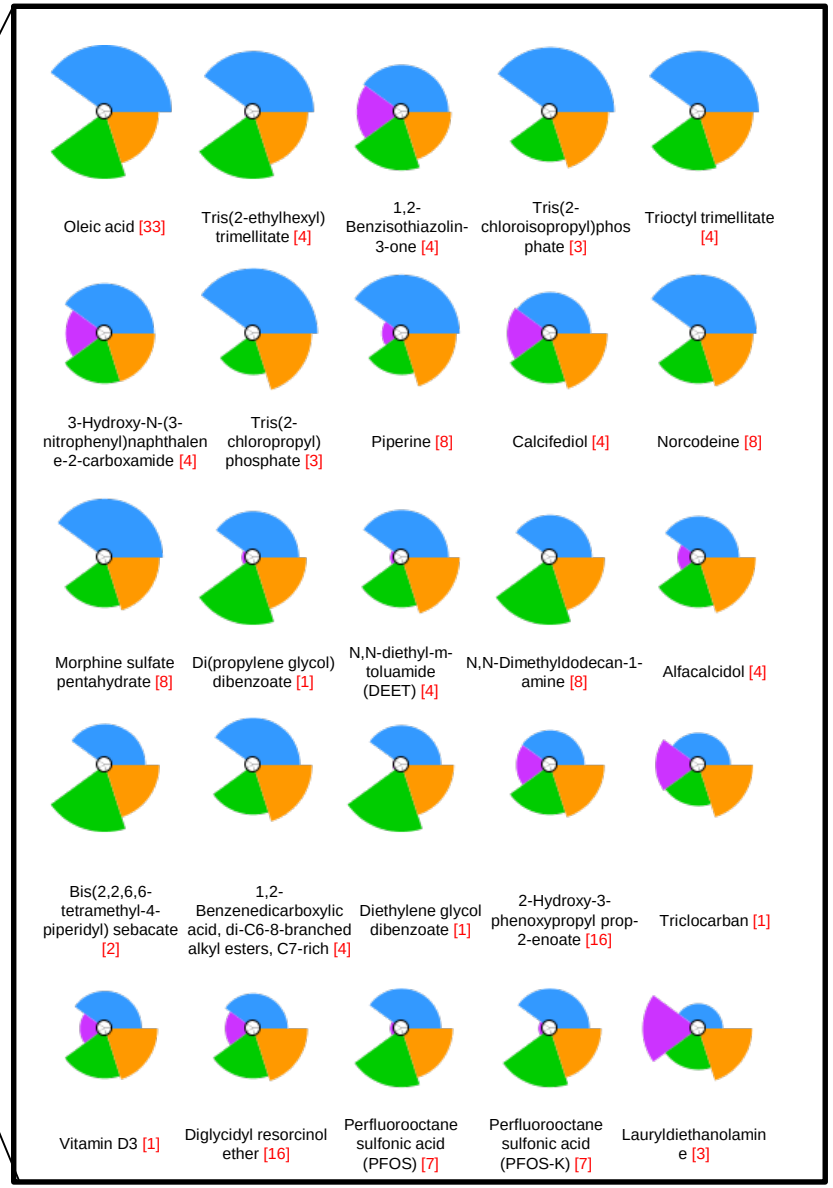
Rager et al. (2016)

Prioritizing Chemical Matches

United States Environmental Protection Agency



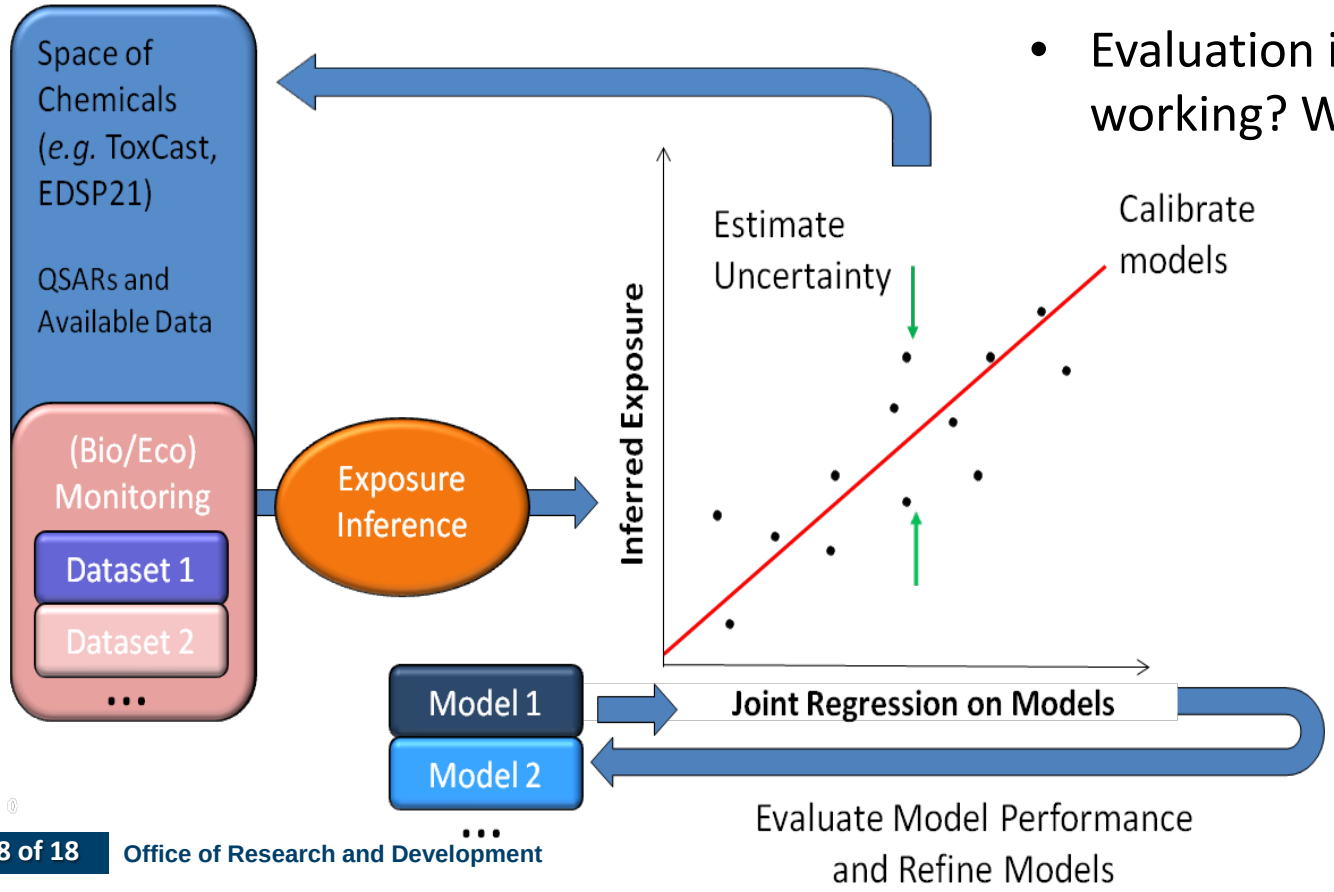




Rager et al. (2016)

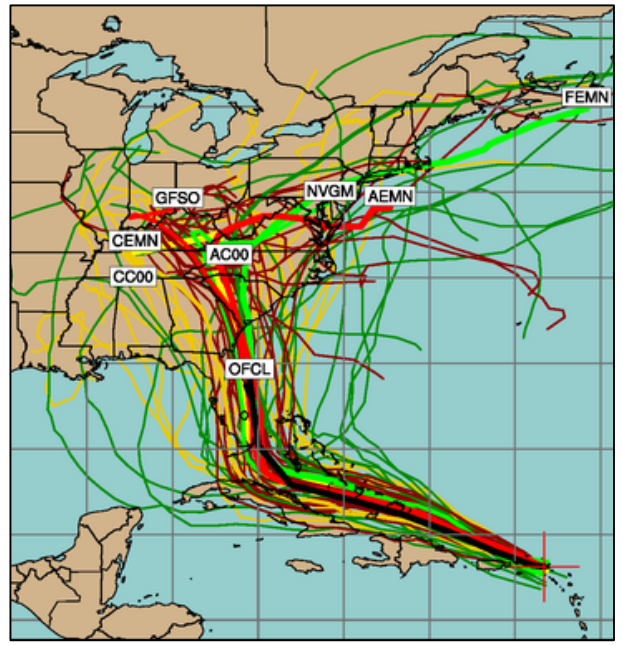
Consensus Exposure Predictions with the SEEM Framework

- Different exposure models incorporate **knowledge**, **assumptions**, and **data** (MacLeod et al., 2010)
- We incorporate multiple models into consensus predictions for 1000s of chemicals within the **Systematic Empirical Evaluation of Models (SEEM)** (Wambaugh et al., 2013, 2014)

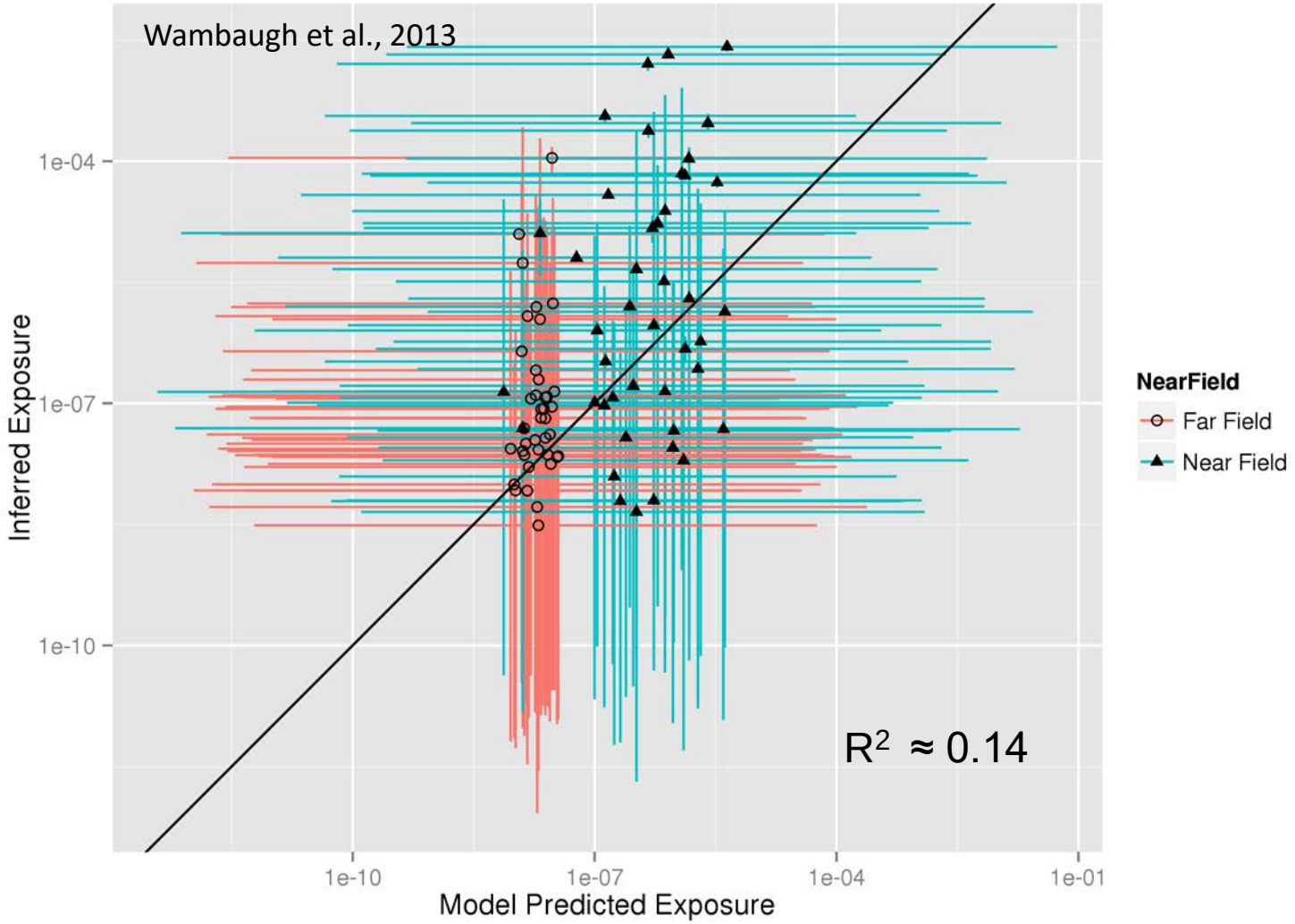


- Evaluation is similar to a sensitivity analysis: What models are working? What data are most needed?

Hurricane Path Prediction is an Example of Integrating Multiple Models

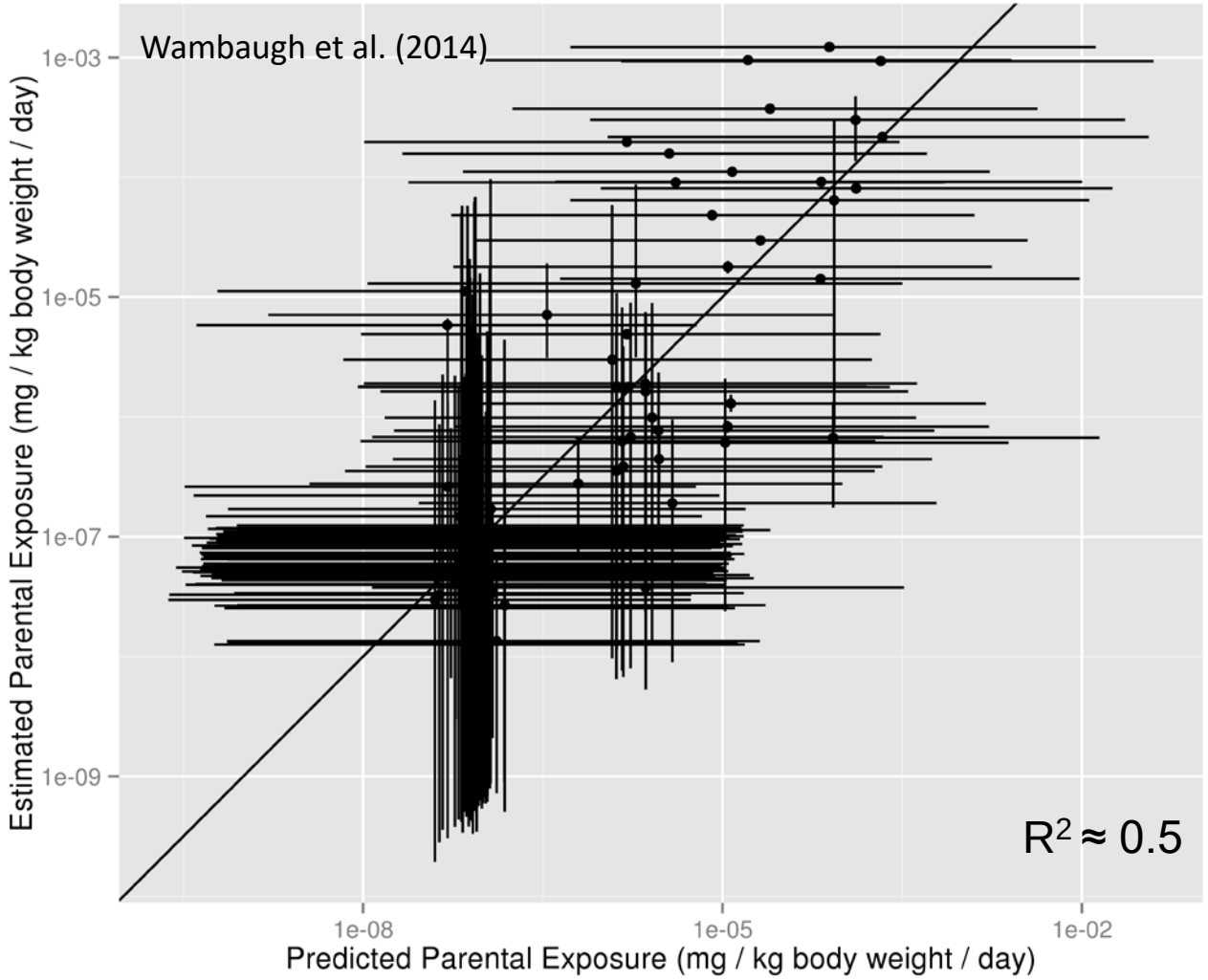


First Generation SEEM



- Those chemicals with "near-field" – proximate, in the home, sources of exposure – had much higher rates of exposure than those with sources outside the home (Wallace et al., 1986)
- The only available "high throughput exposure models in 2013 were for far-field sources

Second Generation SEEM

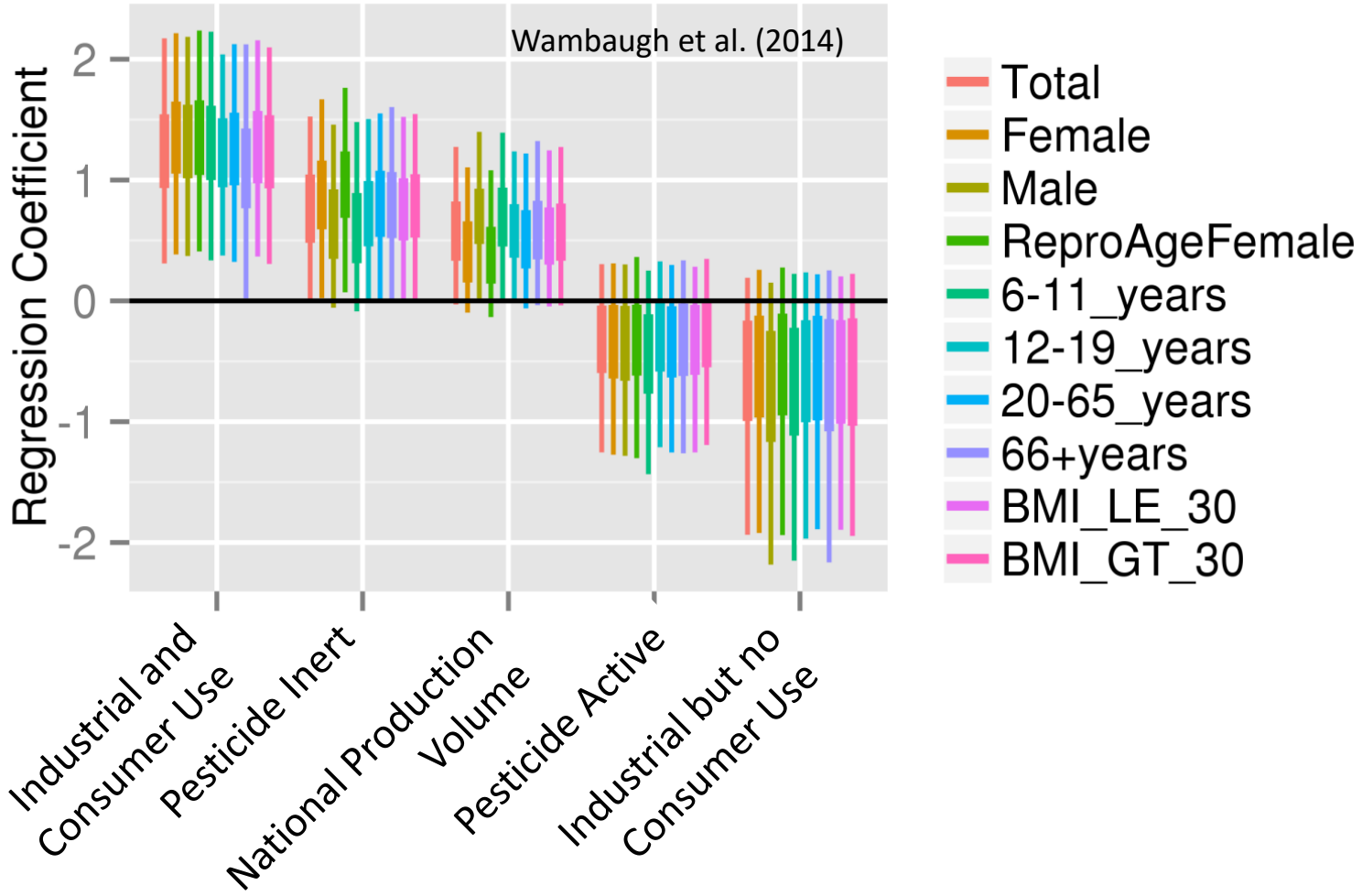


$R^2 \approx 0.5$ indicates that we can predict 50% of the chemical to chemical variability in median NHANES exposure rates

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

Heuristics of Exposure



$R^2 \approx 0.5$ indicates that we can predict 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

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

Predicting Exposure Pathways

We use the method of Random Forests to relate chemical structure and properties to exposure pathway

	NHANES Chemicals	Positives	Negatives	OOB Error Rate	Positives Error Rate	Balanced Accuracy	Sources of Positives	Sources of Negatives
Dietary	24	2523	8865	27	32	73	FDA CEDI, ExpoCast, CPDat (Food, Food Additive, Food Contact), NHANES Curation	Pharmapendium, CPDat (non-food), NHANES Curation
Near-Field	**49**	1622	567	26	24	74	CPDat (consumer_use, building_material), ExpoCast, NHANES Curation	CPDat (Agricultural, Industrial), FDA CEDI, NHANES Curation
Far-Field Pesticide	**94**	1480	6522	21	36	80	REDs, Swiss Pesticides, Stockholm Convention, CPDat (Pesticide), NHANES Curation	Pharmapendium, Industrial Positives, NHANES Curation
Far Field Industrial	**42**	5089	2913	19	16	81	CDR HPV, USGS Water Occurrence, NORMAN PFAS, Stockholm Convention, CPDat (Industrial, Industrial_Fluid), NHANES Curation	Pharmapendium, Pesticide Positives, NHANES Curation

Ring et al., submitted

Ring et al., submitted

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

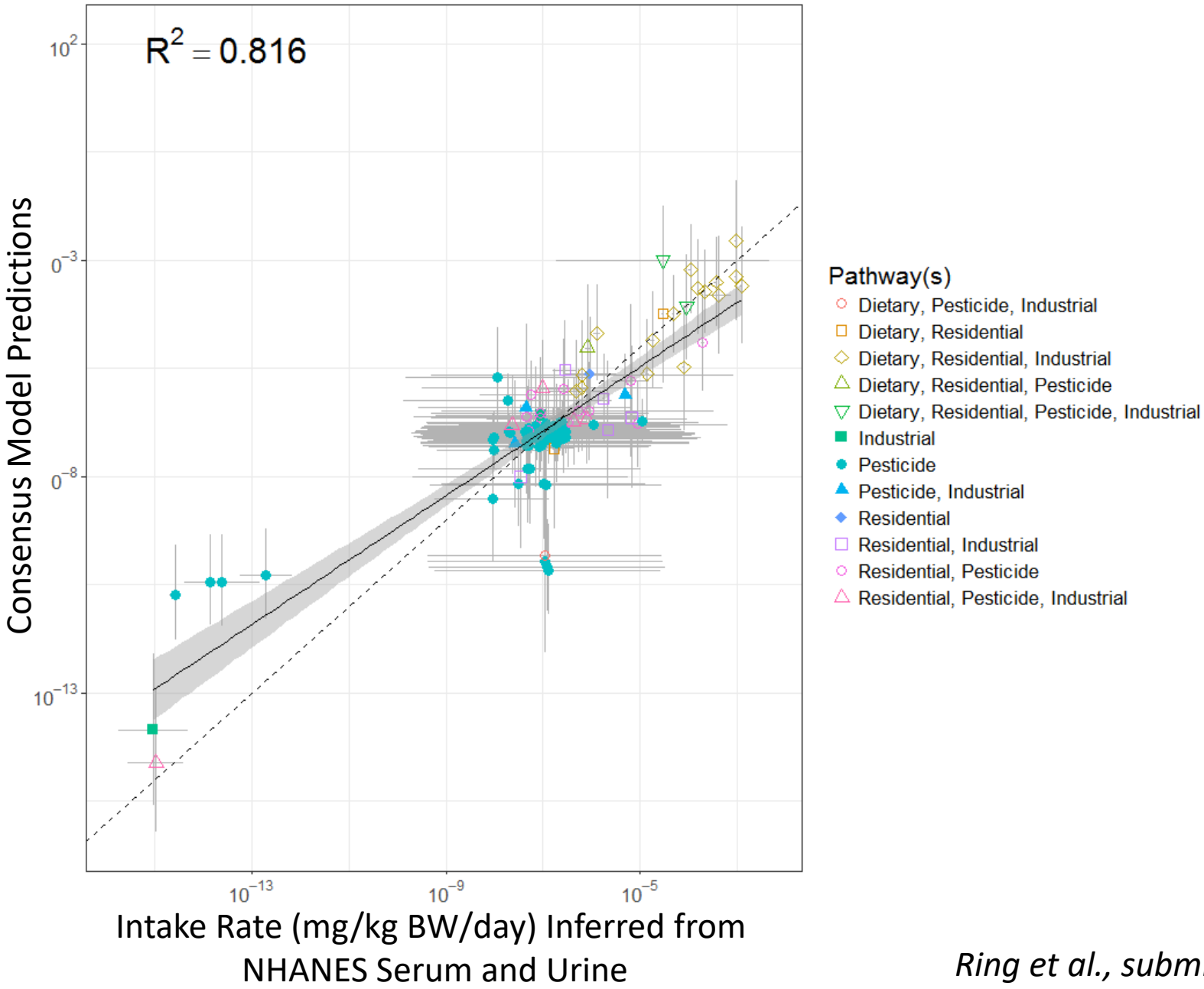
Predictor	Reference(s)	Chemicals Predicted	Pathways
EPA Inventory Update Reporting and Chemical Data Reporting (CDR) (2015)	US EPA (2018)	7856	All
Stockholm Convention of Banned Persistent Organic Pollutants (2017)	Lallas (2001)	248	Far-Field Industrial and Pesticide
EPA Pesticide Reregistration Eligibility Documents (REDs) Exposure Assessments (Through 2015)	Wetmore et al. (2012, 2015)	239	Far-Field Pesticide
United Nations Environment Program and Society for Environmental Toxicology and Chemistry toxicity model (USEtox) Industrial Scenario (2.0)	Rosenbaum et al. (2008)	8167	Far-Field Industrial
USEtox Pesticide Scenario (2.0)	Fantke et al. (2011, 2012, 2016)	940	Far-Field Pesticide
Risk Assessment IDentification And Ranking (RAIDAR) Far-Field (2.02)	Arnot et al. (2008)	8167	Far-Field Pesticide
EPA Stochastic Human Exposure Dose Simulator High Throughput (SHEDS-HT) Near-Field Direct (2017)	Isaacs (2017)	7511	Far-Field Industrial and Pesticide
SHEDS-HT Near-field Indirect (2017)	Isaacs (2017)	1119	Residential
Fugacity-based INdoor Exposure (FINE) (2017)	Bennett et al. (2004), Shin et al. (2012)	645	Residential
RAIDAR-ICE Near-Field (0.803)	Arnot et al., (2014), Zhang et al. (2014)	1221	Residential
USEtox Residential Scenario (2.0)	Jolliet et al. (2015), Huang et al. (2016,2017)	615	Residential
USEtox Dietary Scenario (2.0)	Jolliet et al. (2015), Huang et al. (2016), Ernstoff et al. (2017)	8167	Dietary

Pathway-Based Consensus Modeling of NHANES

- New machine learning tools provide improved high throughput exposure estimates by matching chemicals to exposure pathways and associated calibrated exposure models.

- Exposure predictors (data and models) have been grouped into four pathways (residential, dietary, pesticidal, and industrial) and calibrated via Bayesian multivariate regression using human intake rates inferred for 114 chemicals from a large bio-monitoring survey.

- We have evaluated and calibrated the models using NHANES biomonitoring data



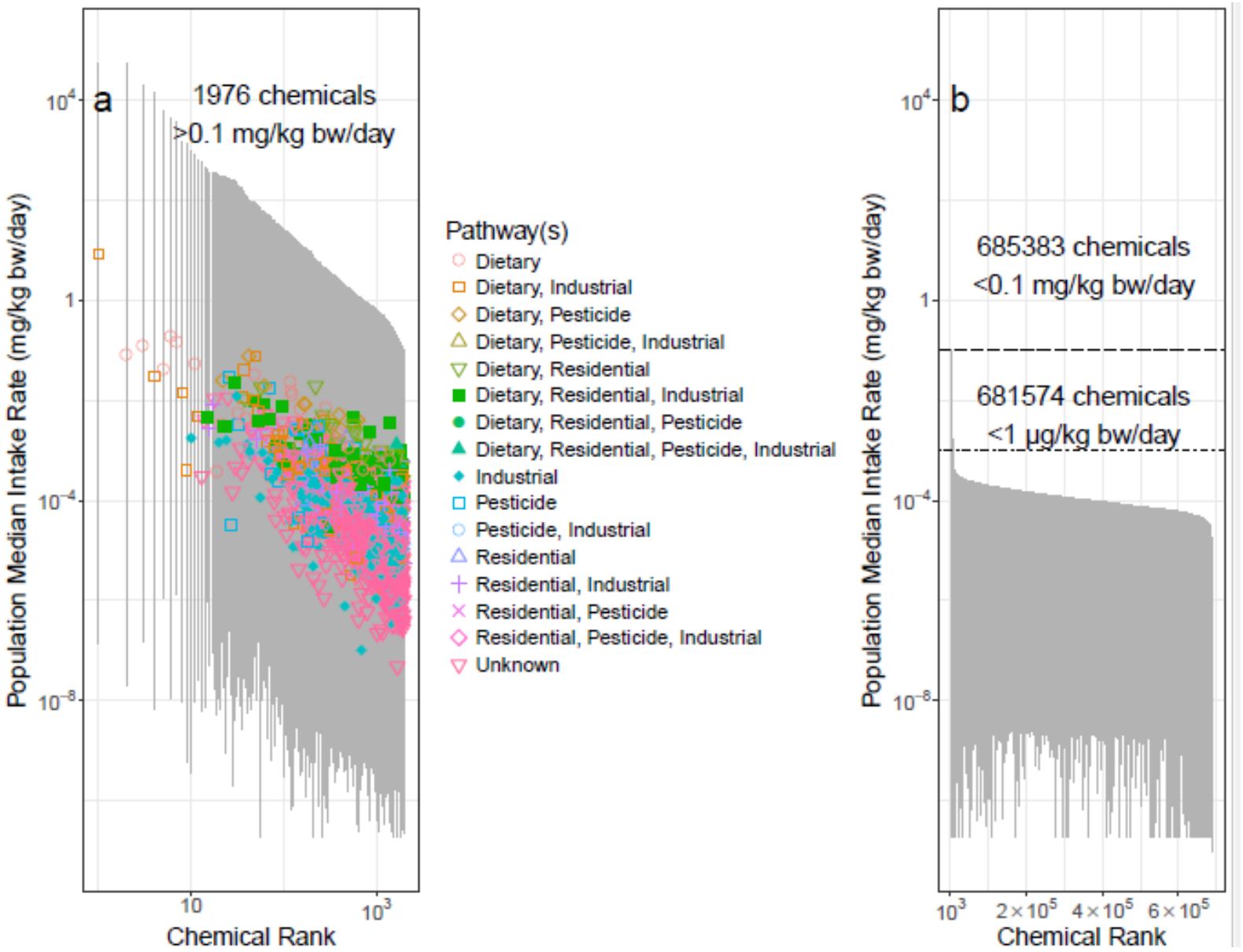
Ring et al., submitted

Consensus Modeling of Median Chemical Intake

- New machine learning tools provide improved high throughput exposure estimates by matching chemicals to exposure pathways and associated calibrated exposure models.

- Exposure predictors (data and models) have been grouped into four pathways (residential, dietary, pesticidal, and industrial) and calibrated via Bayesian multivariate regression using human intake rates inferred for 114 chemicals from a large bio-monitoring survey.

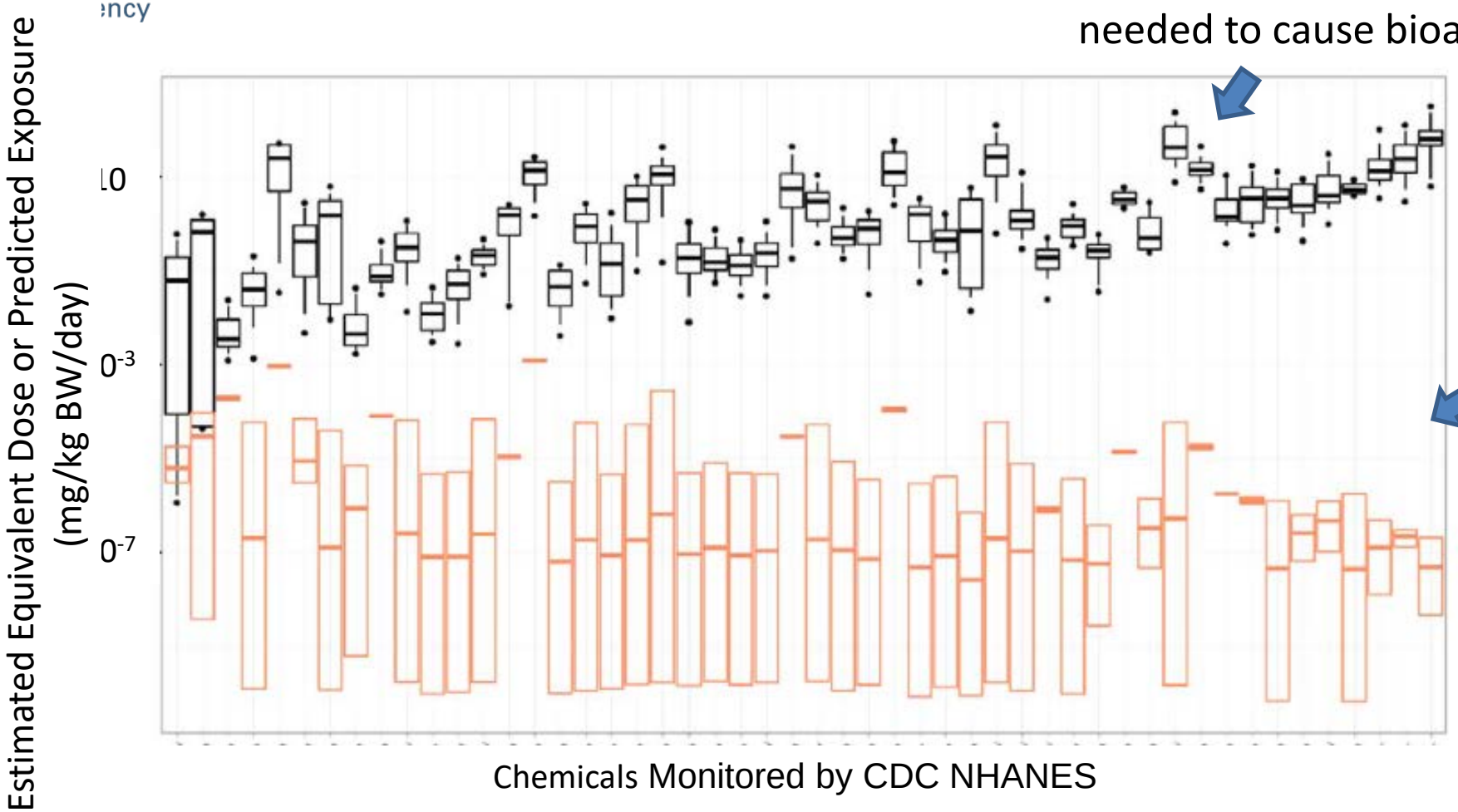
- We have evaluated and calibrated the models using NHANES biomonitoring data



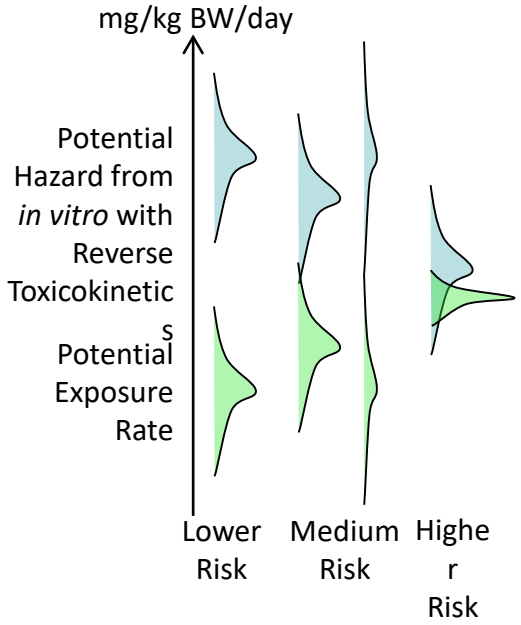
Ring et al., submitted

Selecting Candidates for Prioritization

ToxCast + HTTK can estimate doses needed to cause bioactivity



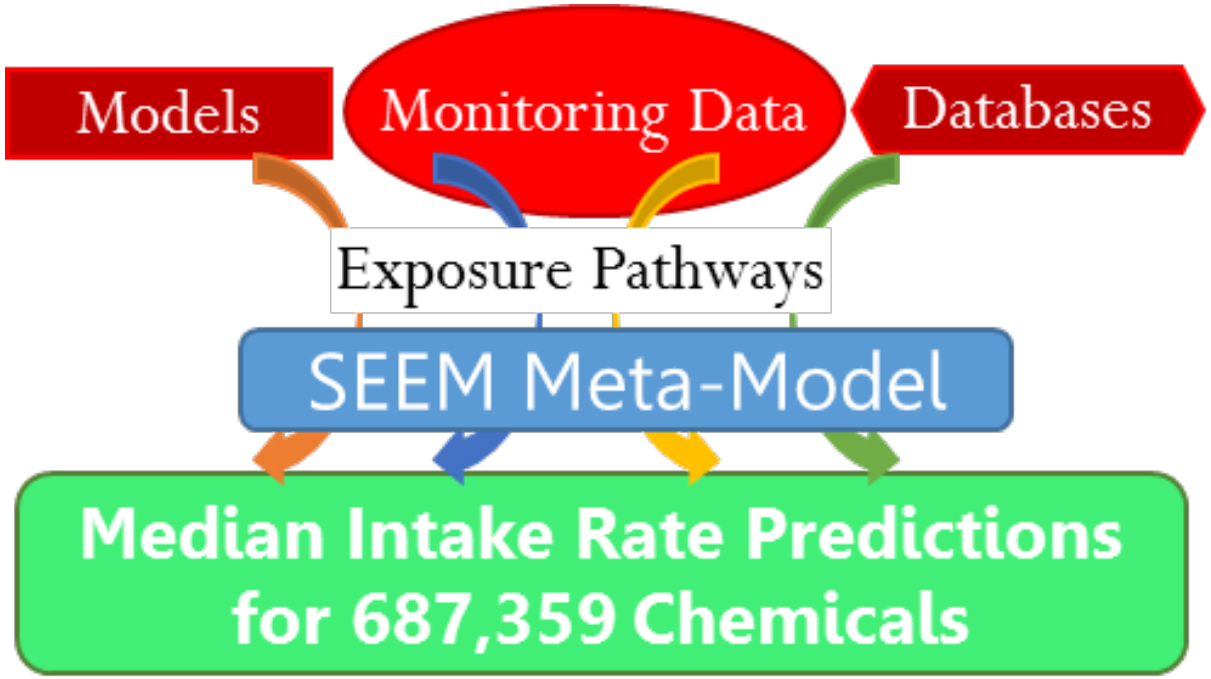
Exposure intake rates can be inferred from biomarkers (Wambaugh et al., 2014)



Ring *et al*. (2017)

Conclusions

- A tapestry of laws covers the chemicals people are exposed to in the United States (Breyer, 2009)
- Most other chemicals, ranging from industrial waste to dyes to packing materials, are covered by the recently updated Toxic Substances Control Act (TSCA) and administered by the EPA
- New approach methodologies (NAMs) are being developed to prioritize these existing and new chemicals for testing
- New machine learning tools provide improved high throughput exposure estimates by matching chemicals to exposure pathways and associated calibrated exposure models.



- Machine learning models based on chemical structure and physico-chemical properties predict whether or not each pathway is relevant to a library of over 680,000 chemicals, allowing an exposure estimate for each chemical based on the calibrated predictors.

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

EPA's Chemical Safety for Sustainability (CSS) Research Program

Rapid Exposure and Dosimetry (RED) Project

NCCT
Chris Grulke
Greg Honda*
Richard Judson
Matthew Linakis*
Andrew McEachran*
Ann Richard
Risa Sayre*
Woody Setzer
Rusty Thomas
John Wambaugh
Antony Williams

NRMRL
Xiaoyu Liu

NHEERL
Linda Adams
Christopher Ecklund
Marina Evans
Mike Hughes
Jane Ellen Simmons

NERL
Cody Addington*
Craig Barber
Namdi Brandon*
Peter Egeghy
Hongtai Huang*
Kristin Isaacs
Ashley Jackson*
Charles Lowe*
Dawn Mills*
Seth Newton
Katherine Phillips
Paul Price
Jeanette Reyes*
Randolph Singh*
Jon Sobus
John Streicher*
Mark Strynar
Mike Tornero-Velez
Elin Ulrich
Dan Vallero
Barbara Wetmore

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

***Trainees**

Collaborators

Arnot Research and Consulting
Jon Arnot
Johnny Westgate
Institut National de l'Environnement et des Risques (INERIS)
Frederic Bois
Integrated Laboratory Systems
Kamel Mansouri
National Toxicology Program
Mike Devito
Steve Ferguson
Nisha Sipes
Ramboll
Harvey Clewell
ScitoVation
Chantel Nicolas
Silent Spring Institute
Robin Dodson
Southwest Research Institute
Alice Yau
Kristin Favela
Summit Toxicology
Lesa Aylward
Technical University of Denmark
Peter Fantke
Tox Strategies
Caroline Ring
Miyoung Yoon
Unilever
Beate Nicol
Cecilie Rendal
Ian Sorrell
United States Air Force
Heather Pangburn
University of California, Davis
Deborah Bennett
University of Michigan
Olivier Jolliet
University of Texas, Arlington
Hyeong-Moo Shin

United States Environmental Protection Agency

References

- Arnot, J. A.; Zhang, X.; Kircanski, I.; Hughes, L.; Armitage, J. Develop Sub-Module for Direct Human Exposures to Consumer Products. Technical Report for the U.S. Environmental Protection Agency; ARC Arnot Research & Consulting, Inc.: Toronto, ON, Canada, 2014.
- Bennett, D. H.; Furtaw, E. J., Fugacity-based indoor residential pesticide fate model. Environmental Science & Technology 2004, 38, (7), 2142-2152.
- 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.
- Ernstoff, A. S.; Fantke, P.; Huang, L.; Jolliet, O., High-throughput migration modelling for estimating exposure to chemicals in food packaging in screening and prioritization tools. Food and Chemical Toxicology 2017, 109, 428-438.
- 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.
- Huang, L.; Ernstoff, A.; Fantke, P.; Csiszar, S. A.; Jolliet, O., A review of models for near-field exposure pathways of chemicals in consumer products. Science of The Total Environment 2017, 574, 1182-1208.
- Huang, L.; Jolliet, O., A parsimonious model for the release of volatile organic compounds (VOCs) encapsulated in products. Atmospheric Environment 2016, 127, 223-235.
- Jolliet, O.; Ernstoff, A. S.; Csiszar, S. A.; Fantke, P., Defining Product Intake Fraction to Quantify and Compare Exposure to Consumer Products. Environmental Science & Technology 2015, 49, (15), 8924-8931.
- Judson, R. S., Magpantay, F. M., Chickarmane, V., Haskell, C., Tania, N., Taylor, J., ... & Houck, K. A. (2015). Integrated model of chemical perturbations of a biological pathway using 18 in vitro high-throughput screening assays for the estrogen receptor. Toxicological Sciences, 148(1), 137-154.
- 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), ORSFl1-ORSFl11.
- 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, H.-M.; McKone, T. E.; Bennett, D. H., Intake Fraction for the Indoor Environment: A Tool for Prioritizing Indoor Chemical Sources. Environmental Science & Technology 2012, 46, (18), 10063-10072.
- 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.
- Zhang, X.; Arnot, J. A.; Wania, F., Model for screening-level assessment of near-field human exposure to neutral organic chemicals released indoors. Environmental science & technology 2014, 48, (20), 12312-12319.