November 29, 2021

Use of computational exposure science & toxicology in chemical risk

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

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

US EPA Office of Research and Development

- The Office of Research and Development (ORD) is the scientific research arm of EPA
  - 555 peer-reviewed journal articles in 2020
- Research is conducted by ORD's four national centers, and three offices organized to address:
  - Public health and env. assessment; comp. tox. and exposure; env. measurement and modeling; and env. solutions and emergency response.
- 13 facilities across the United States
- Research conducted by a combination of Federal scientists (including uniformed members of the **Public Health Service**); 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 chemical signatures in pooled human blood samples, many appear to be exogenous
- CDC National Health and Nutrition Examination Survey (NHANES) monitors biomarkers of hundred for chemicals in the general U.S. population
- A tapestry of laws covers the chemicals people are exposed to in the United States (Breyer, 2009)
- Chemical safety testing is primarily for food additives, pharmaceuticals, and pesticide active ingredients (NRC, 2007)

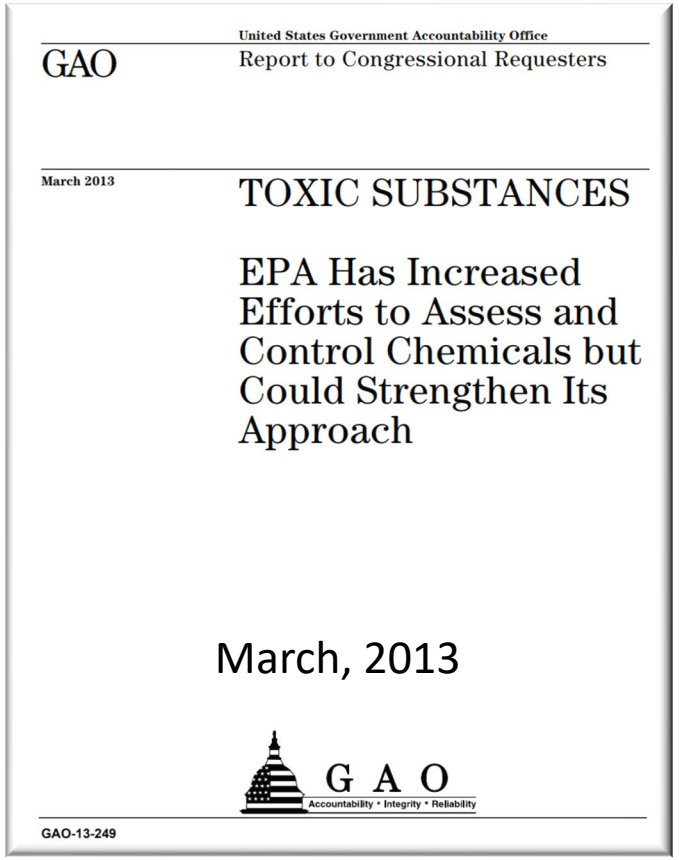
November 29, 2014

Chemical Regulation in the United States

- Chemical safety testing is primarily for food additives, pharmaceuticals, and pesticide active ingredients
  (NRC, 2007)

- Most other chemicals, ranging from industrial waste to dyes to packing materials, are covered by the **Toxic Substances Control Act (TSCA)** administered by the EPA

- Thousands of chemicals on the market were "grandfathered" without assessment
  Judson et al. (2009), Egeghy et al. (2012), Wetmore et al. (2015)


United States Government Accountability Office

GAO

Report to Congressional Requesters

March 2013

TOXIC SUBSTANCES

EPA Has Increased Efforts to Assess and Control Chemicals but Could Strengthen Its Approach

March, 2013

G A O

Accountability • Integrity • Reliability

GAO-13-249

"Tens of thousands of chemicals are listed with the Environmental Protection Agency (EPA) for commercial use in the United States, with an average of 600 new chemicals listed each year."
U.S. Government Accountability Office

Toxic Substances Control Act (TSCA)

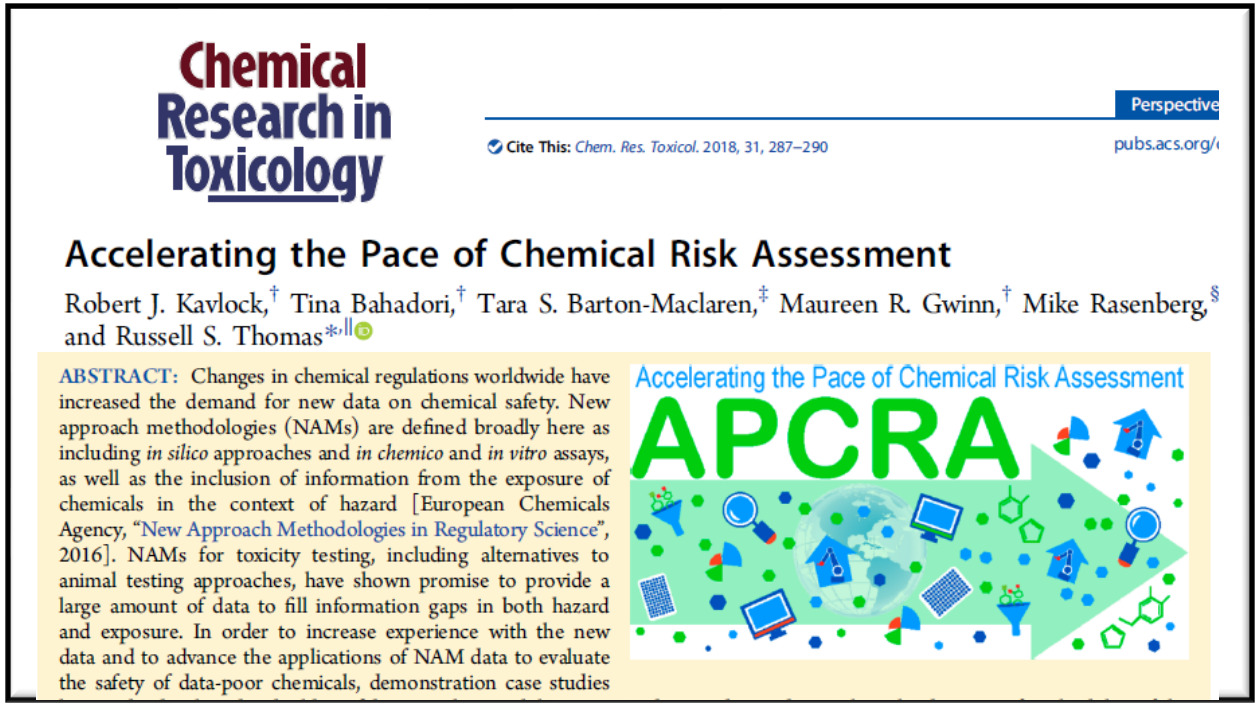
- TSCA was updated in June 2016 to allow more rapid evaluation of chemicals (Frank R. Lautenberg Chemical Safety for the 21st Century Act)
- New approach methodologies (NAMs) are being considered to inform prioritization of chemicals for testing and evaluation (Kavlock et al., 2018)
- EPA has released a "A Working Approach for Identifying Potential Candidate Chemicals for Prioritization" (September, 2018)

Schmidt, C. W. (2016). TSCA 2.0: A new era in chemical risk management", Environmental Health Perspectives, A182-A186.

New Approach Methodologies (NAMs)

- There are roughly 10,000 TSCA-relevant chemicals in commerce
  - Traditional methods are too resource-intensive to address all of these
  - EPA, other US regulators, and international governments are all considering NAMs

- NAMs include:
  - High throughput screening (ToxCast)
  - High throughput exposure estimates (ExpoCast)
  - High throughput toxicokinetics (HTTK)



Chemical Research in Toxicology

Perspective

Cite This: *Chem. Res. Toxicol.* 2018, 31, 287–290

pubs.acs.org/

Accelerating the Pace of Chemical Risk Assessment

Robert J. Kavlock,[†] Tina Bahadori,[†] Tara S. Barton-Maclaren,[‡] Maureen R. Gwinn,[†] Mike Rasenberg,[§] and Russell S. Thomas*[,∥]

ABSTRACT: Changes in chemical regulations worldwide have increased the demand for new data on chemical safety. New approach methodologies (NAMs) are defined broadly here as including *in silico* approaches and *in chemico* and *in vitro* assays, as well as the inclusion of information from the exposure of chemicals in the context of hazard [European Chemicals Agency, "New Approach Methodologies in Regulatory Science", 2016]. NAMs for toxicity testing, including alternatives to animal testing approaches, have shown promise to provide a large amount of data to fill information gaps in both hazard and exposure. In order to increase experience with the new data and to advance the applications of NAM data to evaluate the safety of data-poor chemicals, demonstration case studies

- TSCA Proof of concept (June 2021): Examine ~200 chemicals with ToxCast, ExpoCast, and HTTK
  - HTTK was rate limiter on number of chemicals
  - *"A Proof-of-Concept Case Study Integrating Publicly Available Information to Screen Candidates for Chemical Prioritization under TSCA"* https://cfpub.epa.gov/si/si_public_record_report.cfm?dirEntryID=349776&Lab=CCTE

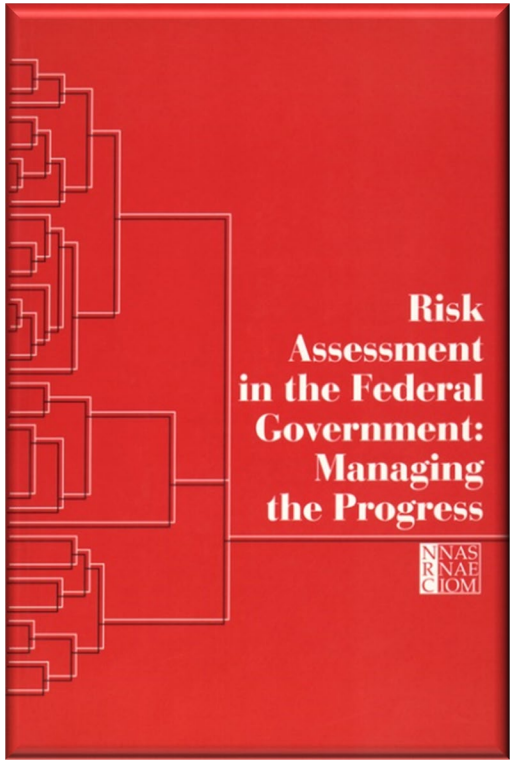
When to Use NAMs?

- …"New approach methods (NAMs), include any technologies, methodologies, approaches or combinations thereof that can be used to provide information on chemical hazard and potential human exposure that can avoid or significantly reduce the use of testing on animals"
  - NAMs for filling information gaps for decision-making
  - integrating data steams into chemical risk assessment
  - making the information publicly available

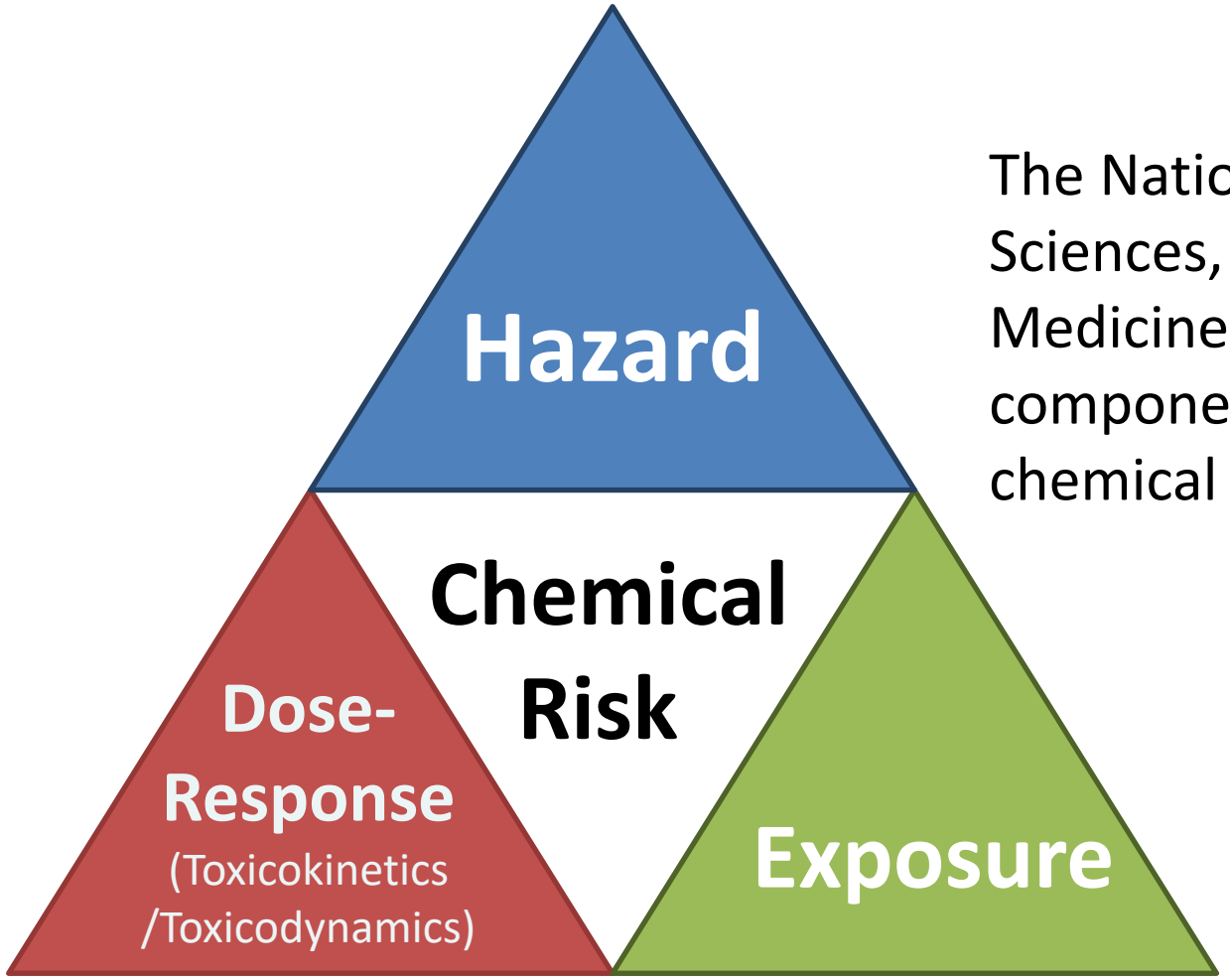
- Can ask (at least) two different questions:
  - For chemicals where we have resources to study:
    When are NAMs equivalent to or better than existing methodologies?
    - Animal testing
    - Clinical studies (for therapeutics)
    - Biomonitoring / epidemiology

  - For chemicals where there are limited resources:
    When are NAMs sufficient to act upon in the absence of other information?
    - Raising/lowering chemical priority?

Three Components for Chemical Risk



NRC (1983)

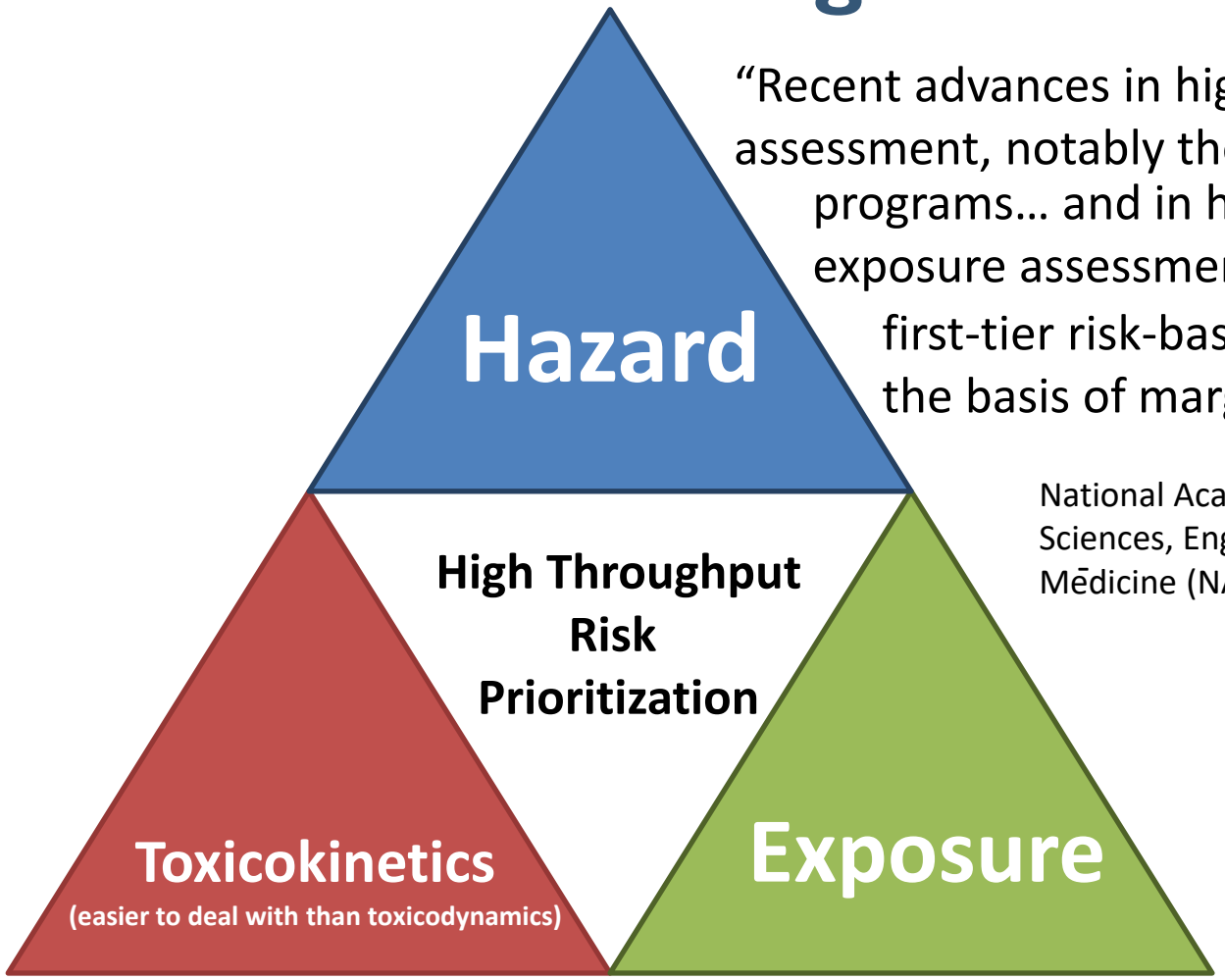


The National Academy of Sciences, Engineering and Medicine (1983) outlined three components for determining chemical risk.

EPA United States Environmental Protection Agency

Calculating Chemical Risk

THE NATIONAL ACADEMIES PRESS
Washington, DC
www.nap.edu
January 5, 2017



"Recent advances in high throughput toxicity assessment, notably the **ToxCast and Tox21** programs… and in high throughput computational exposure assessment [**ExpoCast**] have enabled first-tier risk-based rankings of chemicals on the basis of margins of exposure"

National Academies of Sciences, Engineering, and Medicine (NASEM)

ToxCast = Toxicity Forecaster
ExpoCast = Exposure Forecaster

New approach methodologies (NAMs) enable risk assessors to more rapidly address public health challenges and chemical regulation

Risk Assessment in the 21st Century

"…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**."

THE NATIONAL ACADEMIES PRESS
Washington, DC
www.nap.edu
January 5, 2017

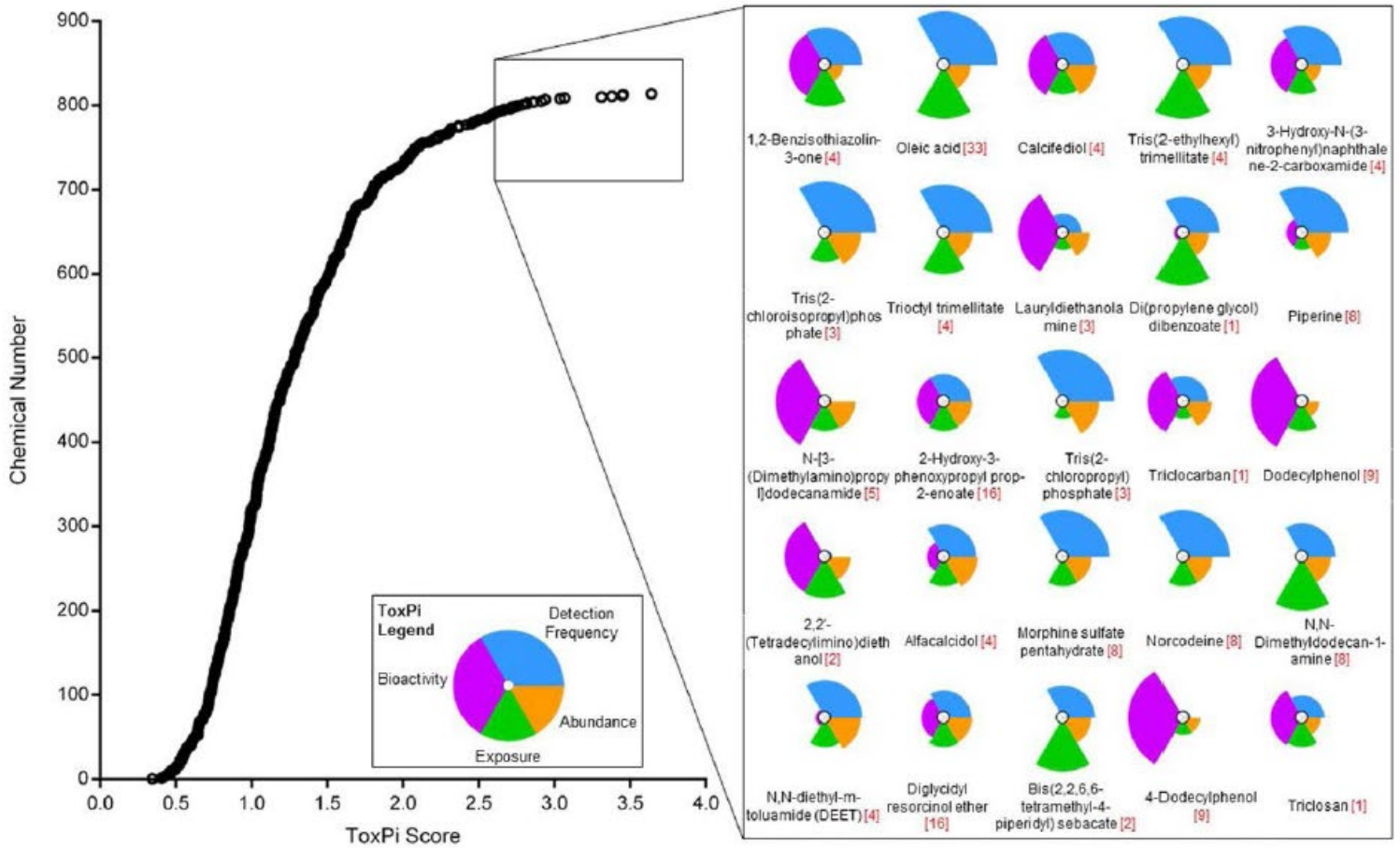
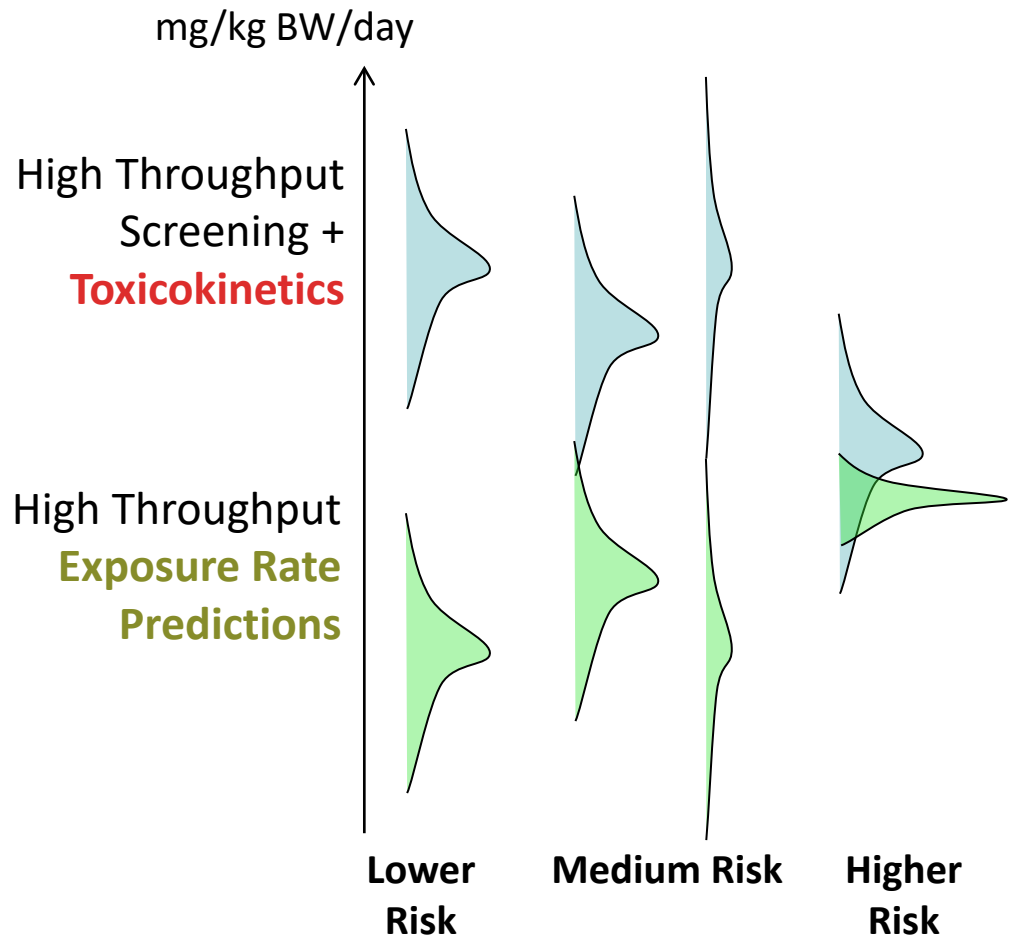


FIGURE 2-7 Data from nontargeted and targeted analysis of dust samples were used with toxicity data to rank chemicals for further analysis and testing. Source: Rager et al. 2016

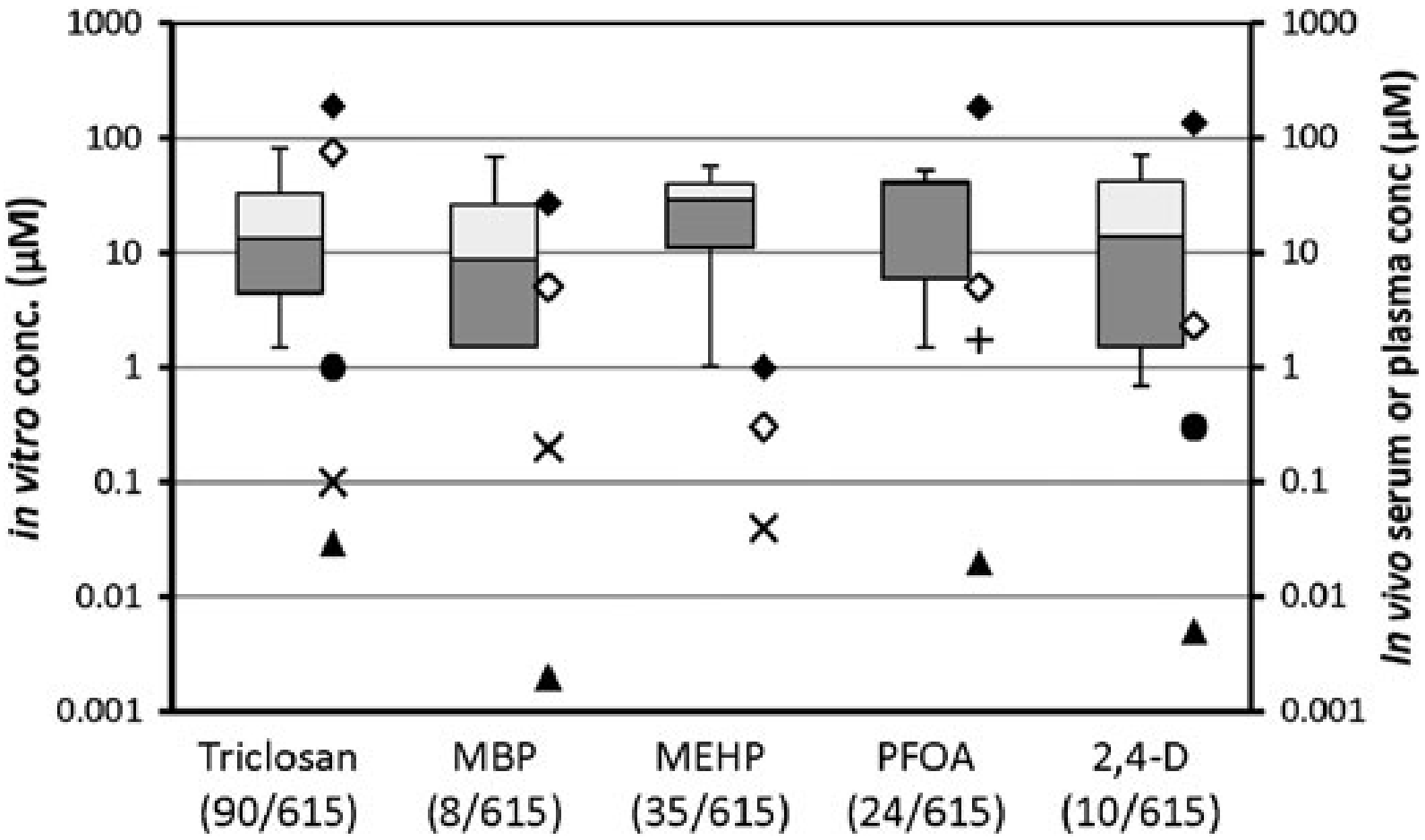
Chemical Risk = Hazard x Exposure

- The U.S. National Research Council (1983) identified chemical risk as a function of both inherent hazard and exposure

- Therefore, high throughput risk prioritization needs:
  1. High throughput hazard characterization (Dix et al., 2007, Collins et al., 2008)
  2. High throughput exposure forecasts (Wambaugh et al., 2013, 2014)
  3. High throughput toxicokinetics (that is, dose-response relationship) linking hazard and exposure (Wetmore et al., 2012, 2015)



The Margin Between Exposure and Hazard



Range of bioactive concentrations across ToxCast assays

Estimated or measured average concentrations associated with the LOAEL in animal studies

NOAEL in animal studies

Humans with chronic exposure reference values (solid circles)

X Volunteers using products containing the chemical

\+ Bio-monitored occupational populations

General populations

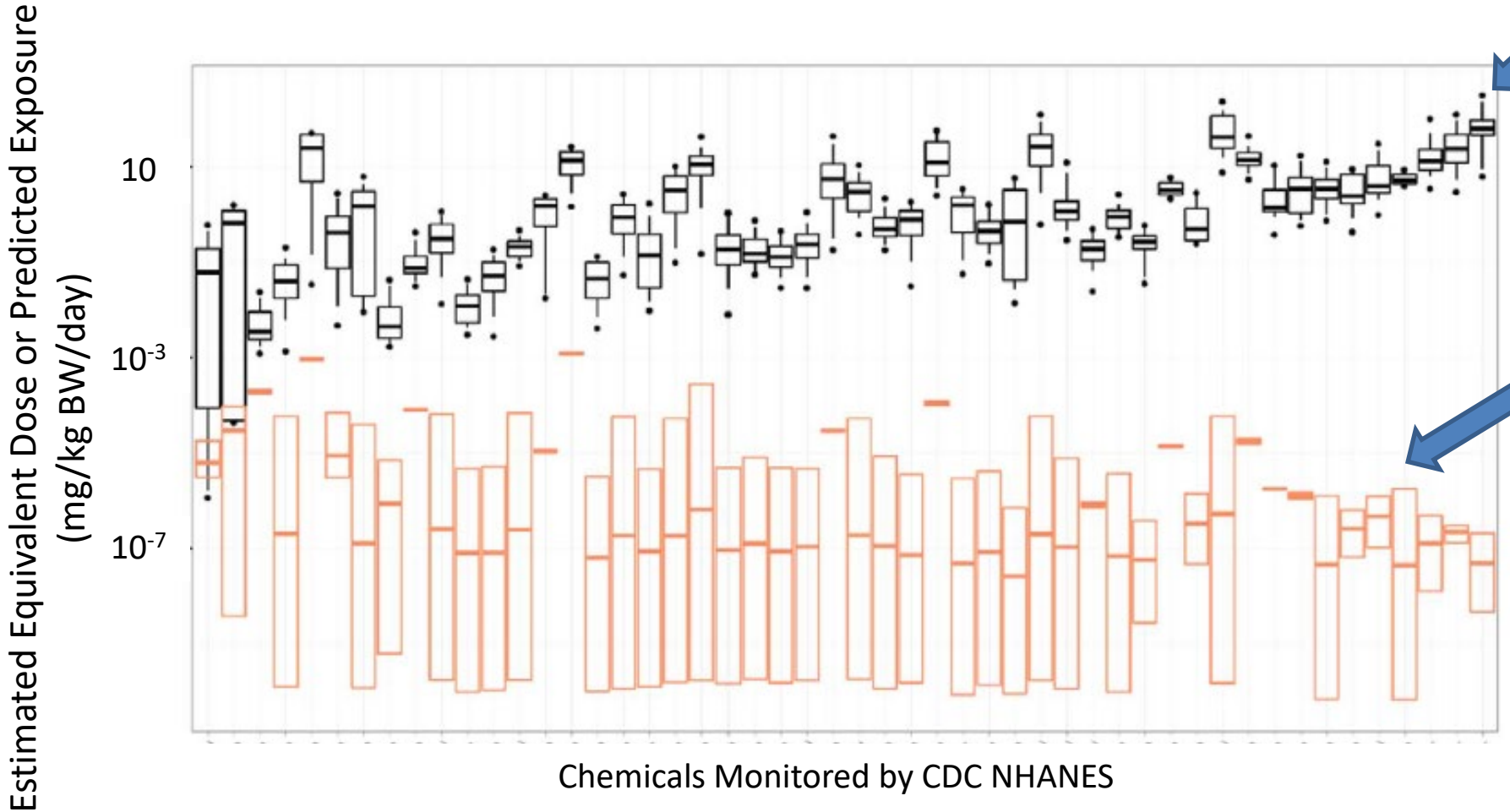
The five chemicals (as of 2011) with plasma biomonitoring AND ToxCast data… what do we do about the other 1000's?

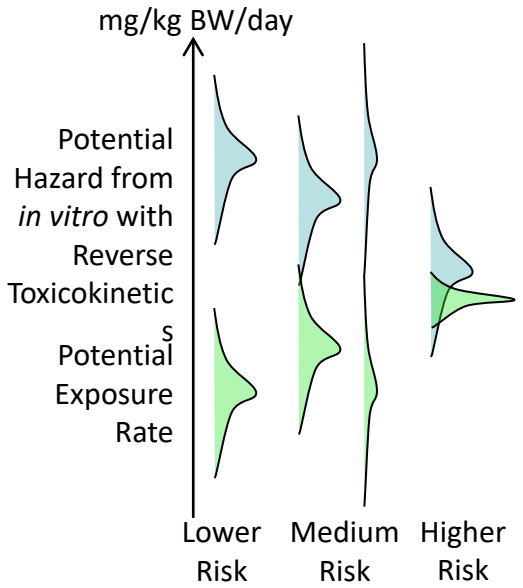
Aylward and Hays (2011)

Chemical Prioritization NAMs

High throughput *in vitro* screening can estimate doses needed to cause bioactivity
(for example, Wetmore et al., 2015)

Exposure intake rates can be inferred from biomarkers
(for example, Ring et al., 2018)

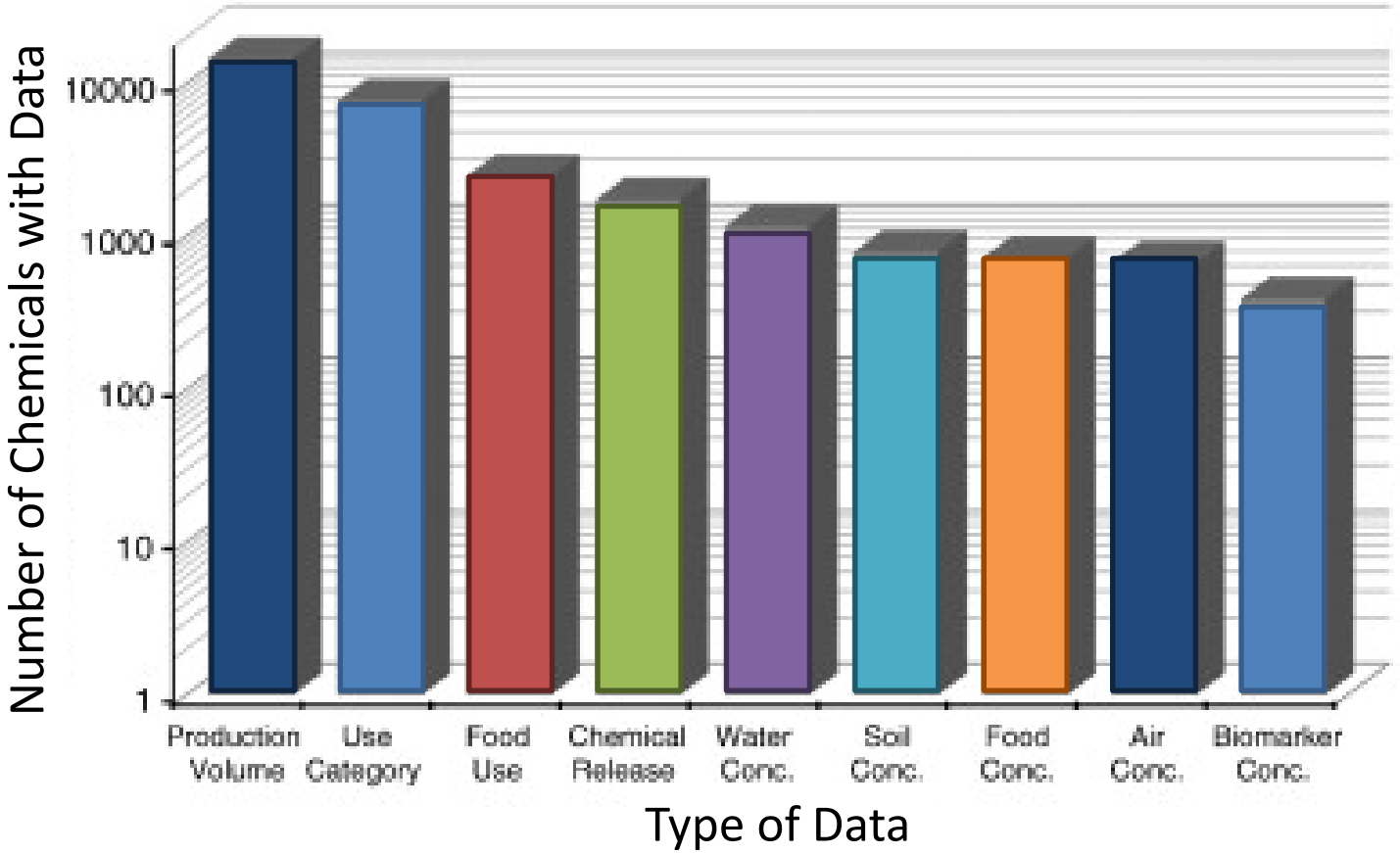




Ring *et al*. (2017)

There are Limited Available Data for Exposure Estimation

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



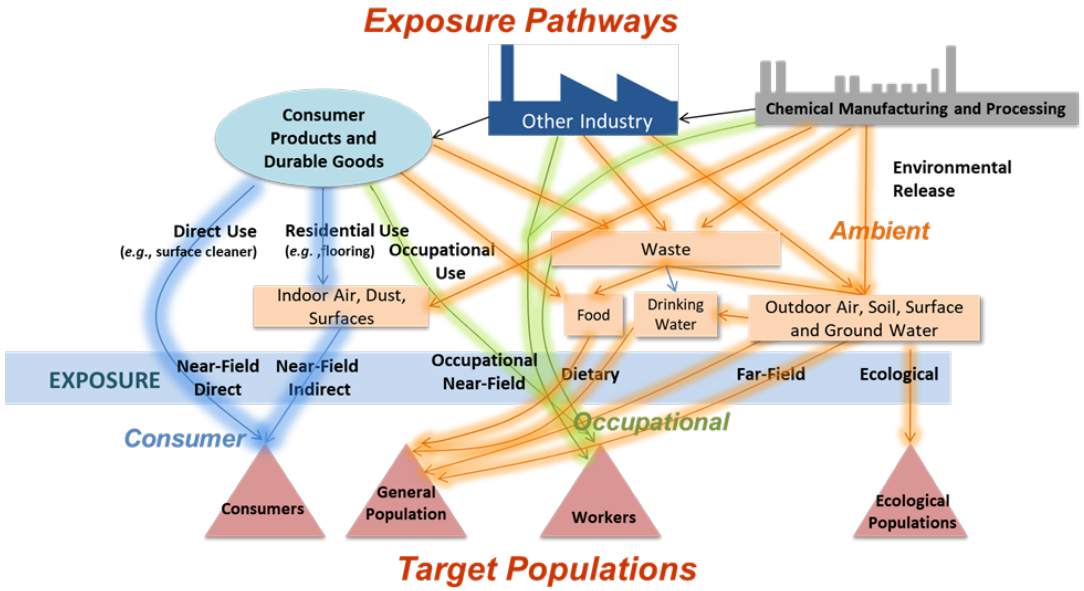
Can we develop new tools to generate the exposure information we need?

NAMs for Exposure Science

- There are at least 10,000 chemicals produced, used in commerce, and potentially present in the environment
  - Traditional methods are too resource-intensive to address all of these
  - New Approach Methodologies (NAMs) have the potential to address these gaps

- The tools to characterize both toxicity and exposure have evolved significantly in the past decade

- NAMs for exposure science are being developed to enable risk assessors to more rapidly address public health challenges and chemical regulation



Exposure is a Complex System

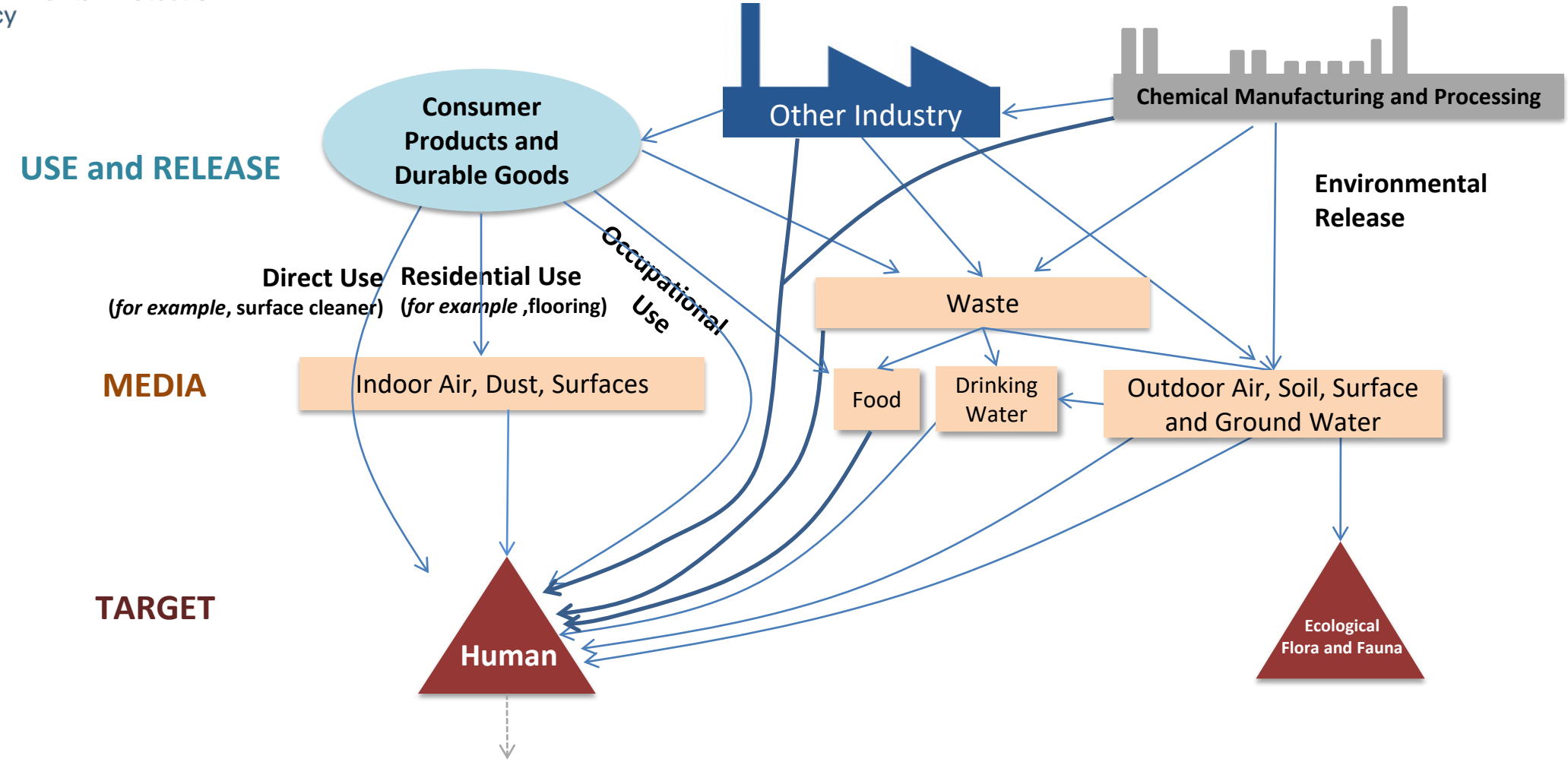
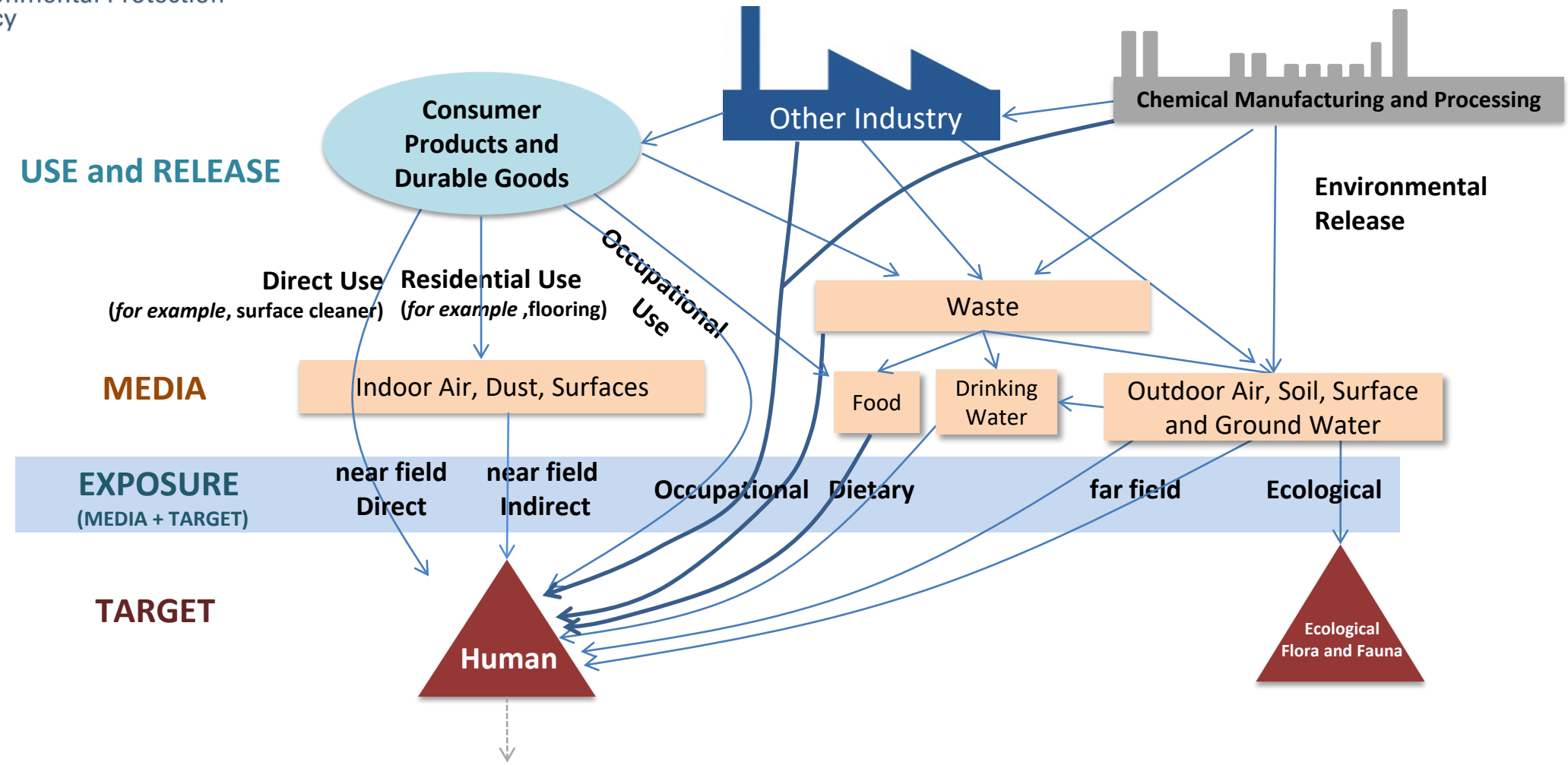


Figure from Kristin Isaacs

The Exposure Event is Often 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)

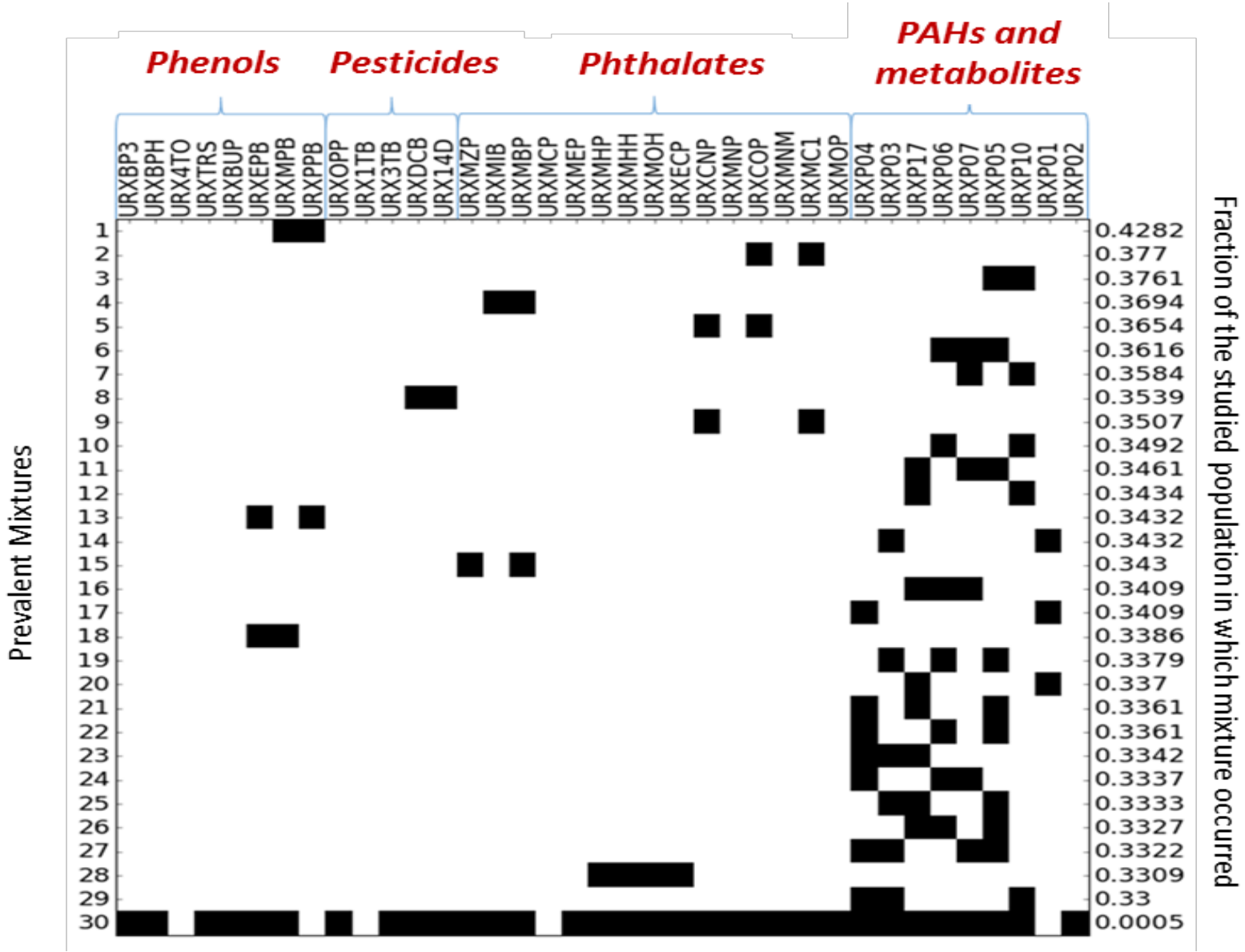
Figure from Kristin Isaacs

What Do We Know About Exposure? Biomonitoring Data

- Centers for Disease Control and Prevention (CDC) National Health and Nutrition Examination Survey (NHANES) provides an important tool for monitoring public health
- Large, ongoing CDC survey of US population: demographic, body measures, medical exam, biomonitoring (health and exposure), …
- Designed to be representative of US population according to census data
- Data sets publicly available (http://www.cdc.gov/nchs/nhanes.htm)
- Includes measurements of:
  - Body weight
  - Height
  - **Chemical analysis of blood and urine**

Identifying Prevalent Mixtures in the NHANES Data

- We used data-mining methods (frequent itemset mining or FIM, Borgelt, 2012) to identify combinations of items (chemicals) that co-occur together within samples from same individual
- Identified a few dozen mixtures present in >30% of U.S. population



Kapraun *et al*. (2017)

What Do We Know About Exposure? Exposure Models

- Any model, including those for exposure, capture knowledge and a hypothesis of how the world works
- EPA's EXPOsure toolBOX (EPA ExpoBox) is a toolbox created to assist individuals from within government, industry, academia, and the general public with assessing exposure
  - Includes many, many models (https://www.epa.gov/expobox)
- These models can be coarsely grouped (Arnot *et al.*, 2006) into:
  - Models that describe "**near field"** sources that are close to the exposed individual (consumer or occupational exposures)
  - Models that describe "**far field"** scenarios wherein individuals are exposed to chemicals that were released or used far away (ambient exposure)

Everyone Uses Models

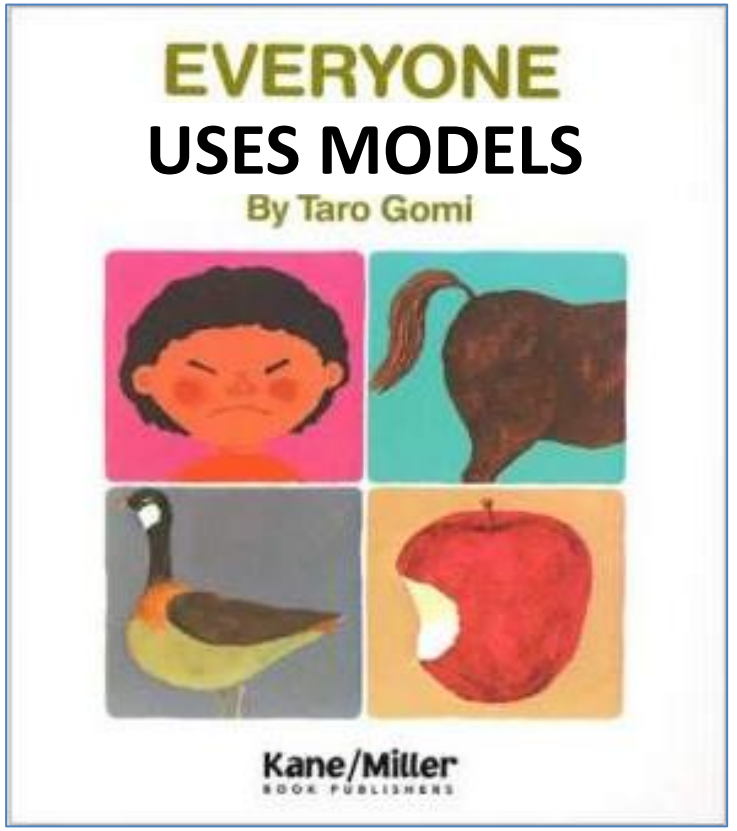
- Toxicology has long relied upon model animal species

- People rely on mental models every day
  - For example, repetitive activities like driving home from work

- Mathematical models offer some significant advantages:
  - Reproducible
  - Can (and should) be transparent

- …with some disadvantages:
  - Sometimes reality is complex
  - Sometimes the model doesn't always work well
  - How do we know we can extrapolate?

- …that can be turned into advantages:
  - If we have evaluated confidence/uncertainty and know the "domain of applicability" we can make better use of mathematical models



Fit for Purpose Models

- A “fit for purpose” model is an abstraction of a complicated problem that allows us to reach a decision.

  “Now it would be very remarkable if any system existing in the real world could be *exactly* represented by any simple model. However, cunningly chosen **parsimonious models** often do provide remarkably useful approximations… **The only question of interest is ‘Is the model illuminating and useful?’**”
  George Box

- A fit for purpose model is **defined as much by what is omitted as what is included** in the model.

- We must accept that there will always be areas in need of better data and models – our knowledge will always be incomplete, and thus we wish to extrapolate.

  - How do I drive to a place I’ve never been before?

NAMs for Exposure Science

Exposure NAM Class	Description	Traditional Approach	NAM Makes Use of					
			Measurement	Toxicokinetics	Models	Descriptors	Evaluation	Machine Learning
Measurements	New techniques including screening analyses capable of detecting hundreds of chemicals present in a sample	Targeted (chemical-specific) analyses	-	●	●	●		●
Toxicokinetics	High throughput methods using *in vitro* data to generate chemical-specific models	Analyses based on *in vivo* animal studies	●	-		●		●
HTE Models	Models capable of making predictions for thousands of chemicals	Models requiring detailed, chemical- and scenario-specific information	●	●	-	●		
Chemical Descriptors	Informatic approaches for organizing chemical information in a machine-readable format	Tools targeted at single chemical analyses by humans				-		●
Evaluation	Statistical approaches that use the data from many chemicals to estimate the uncertainty in a prediction for a new chemical	Comparison of model predictions to data on a per chemical basis	●	●	●	●	-	●
Machine Learning	Computer algorithms to identify patterns	Manual Inspection of the Data	●	●		●		-
Prioritization	Integration of exposure and other NAMs to identify chemicals for follow-up study	Expert decision making	●	●	●	●	●	●

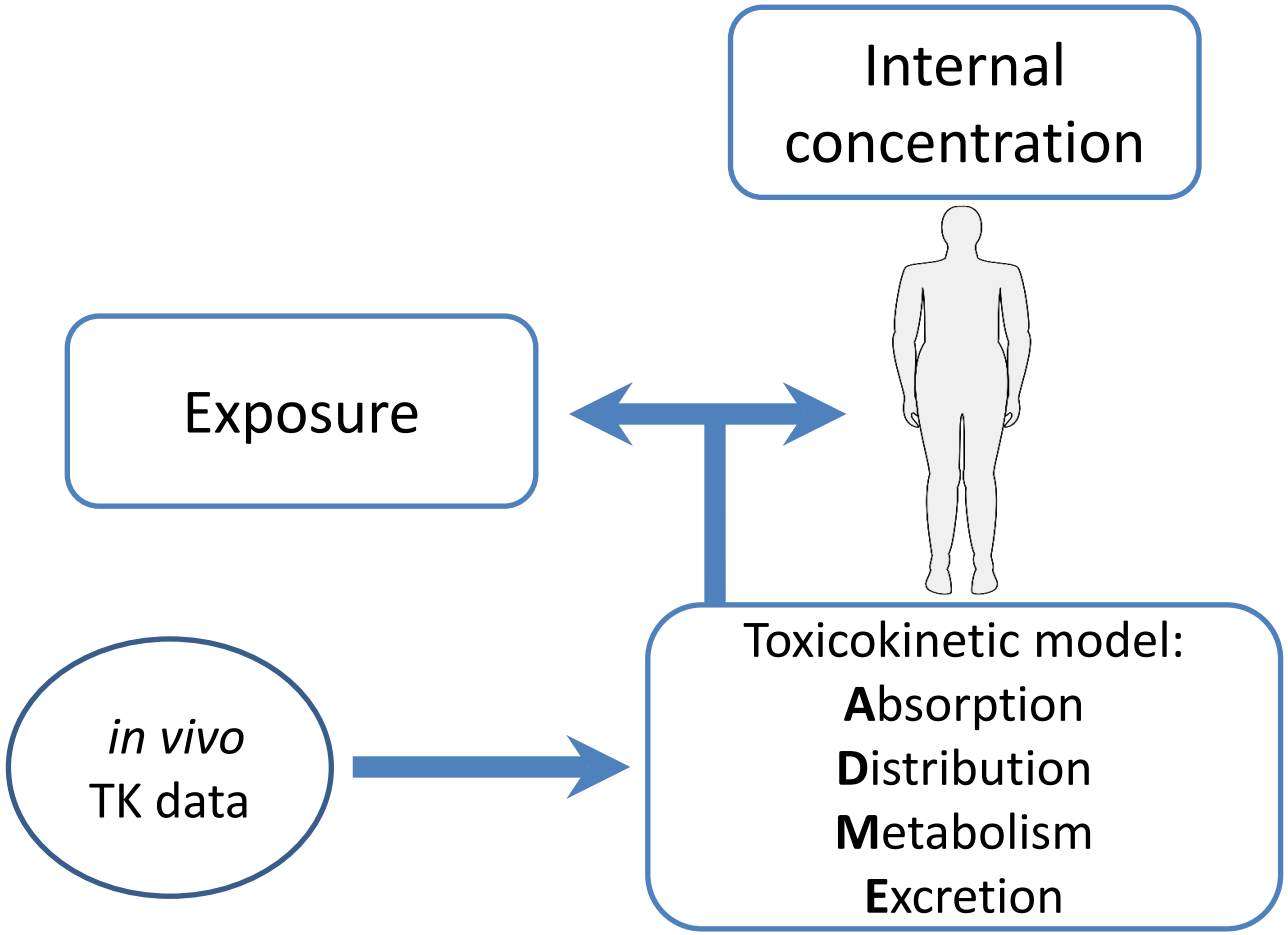
Wambaugh et al., (2019)

NAMs for Exposure Science

Exposure NAM Class	Description	Traditional Approach	NAM Makes Use of					
			Measurement	Toxicokinetics	Models	Descriptors	Evaluation	Machine Learning
Measurements	New techniques including screening analyses capable of detecting hundreds of chemicals present in a sample	Targeted (chemical-specific) analyses	-	●	●	●		●
Toxicokinetics	High throughput methods using *in vitro* data to generate chemical-specific models	Analyses based on *in vivo* animal studies	●	-		●		●
HTE Models	Models capable of making predictions for thousands of chemicals	Models requiring detailed, chemical- and scenario-specific information	●	●	-	●		
Chemical Descriptors	Informatic approaches for organizing chemical information in a machine-readable format	Tools targeted at single chemical analyses by humans				-		●
Evaluation	Statistical approaches that use the data from many chemicals to estimate the uncertainty in a prediction for a new chemical	Comparison of model predictions to data on a per chemical basis	●	●	●	●	-	●
Machine Learning	Computer algorithms to identify patterns	Manual Inspection of the Data	●	●		●		-
Prioritization	Integration of exposure and other NAMs to identify chemicals for follow-up study	Expert decision making	●	●	●	●	●	●

Wambaugh et al., (2019)

Toxicokinetics



- Toxicokinetics describes the absorption, distribution, metabolism, and excretion of a chemical by the body:
  - Chemical-specific
  - Links exposure with internal concentrations

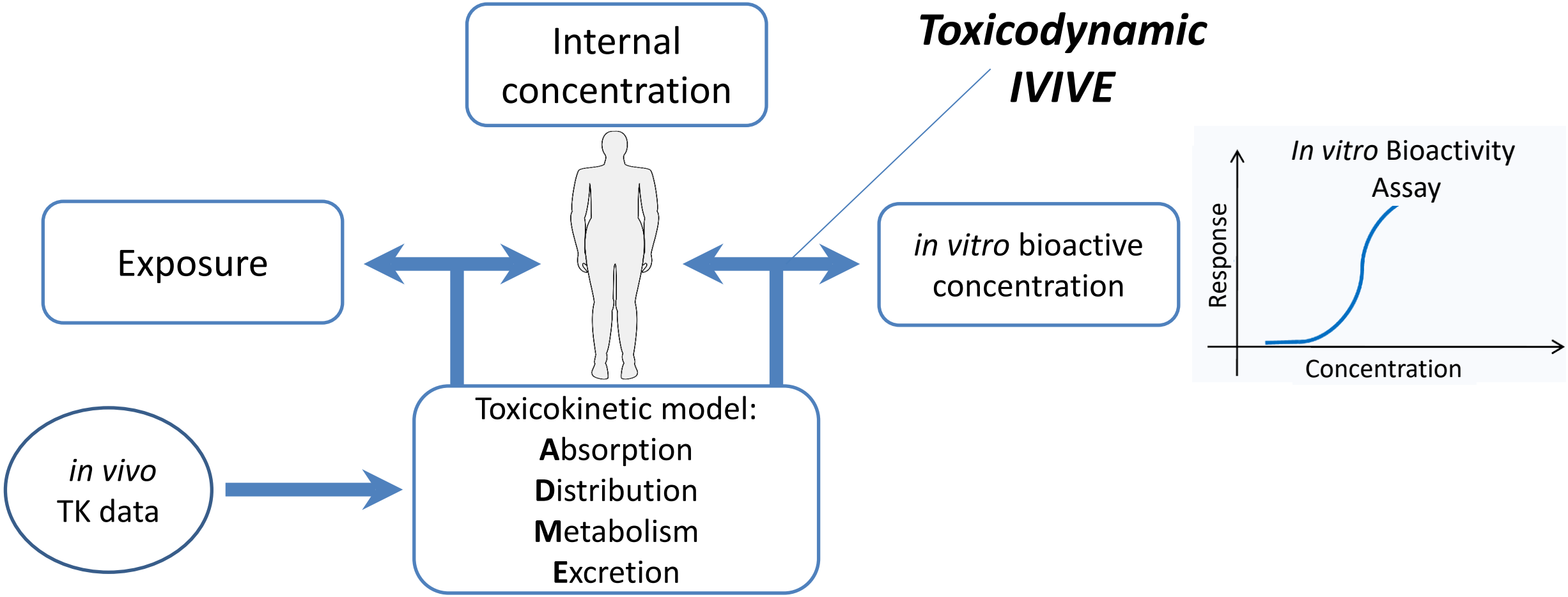
Breen et al. (submitted)

HTTK: A NAM for Exposure

- To provide toxicokinetic data for larger numbers of chemicals collect *in vitro*, high throughput toxicokinetic (HTTK) data (for example, Rotroff et al., 2010, Wetmore et al., 2012, 2015)
- HTTK methods have been used by the pharmaceutical industry to determine range of efficacious doses and to prospectively evaluate success of planned clinical trials (Jamei, *et al.*, 2009; Wang, 2010)
- The **primary goal** of HTTK is to provide a human dose context for bioactive *in vitro* concentrations from HTS (that is*, in vitro-in vivo* extrapolation, or **IVIVE**) (for example, Wetmore et al., 2015)
- A **secondary goal** is to provide **open-source data and models** for evaluation and use by the broader scientific community (Pearce et al, 2017a)

In Vitro-In Vivo Extrapolation (IVIVE)

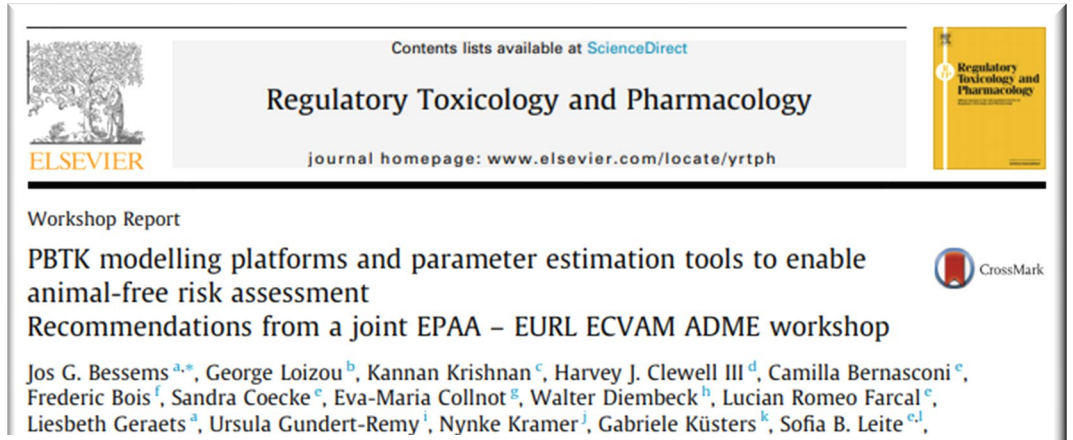
- Translation of *in vitro* high throughput screening requires chemical-specific toxicokinetic models
  - Needed for anywhere from dozens to thousands of chemicals

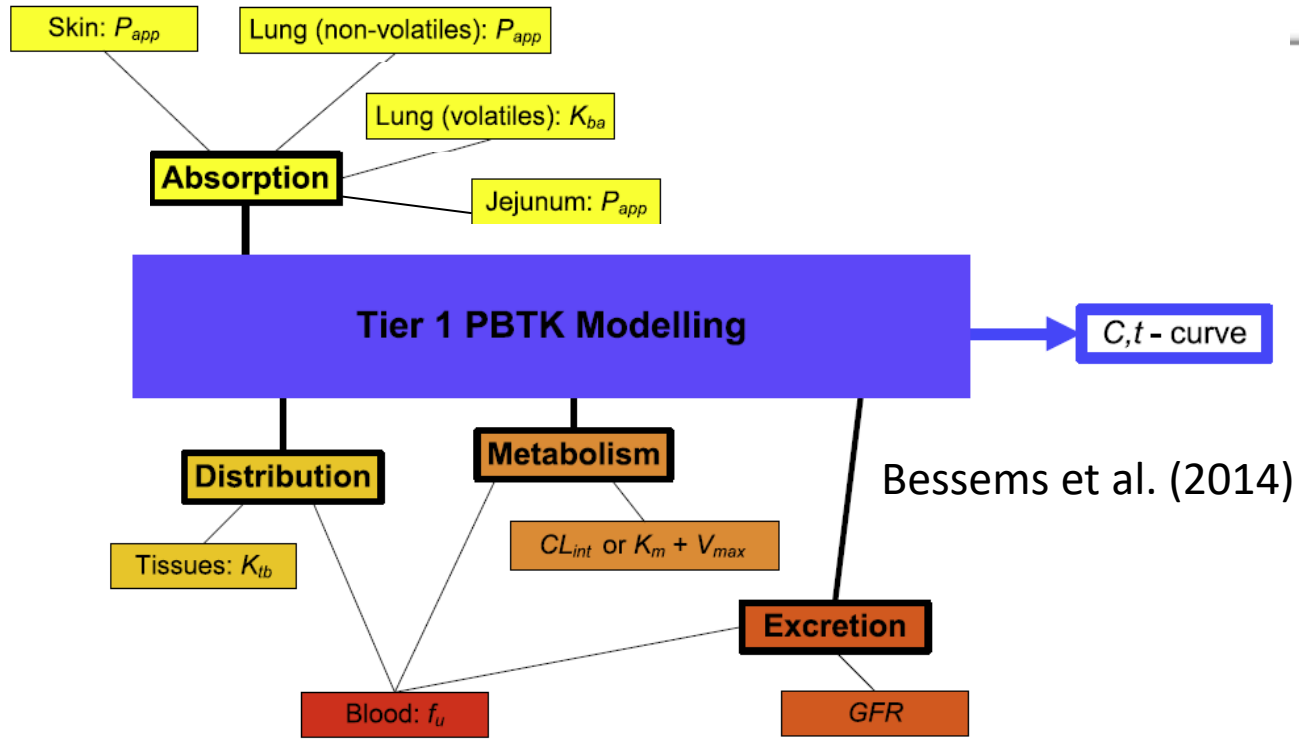


Breen et al. (submitted)

Fit for Purpose IVIVE

- We choose to make the complexity of the model and the number of physiological processes appropriate to decision context
- Bessems et al. (2014): We need “a first, relatively quick (‘Tier 1’), estimate” of concentration vs. time in blood, plasma, or cell
- They suggested that we neglect active metabolism – thanks to *in vitro* measurements we can now do better
- We do neglect transport and other protein-specific phenomena

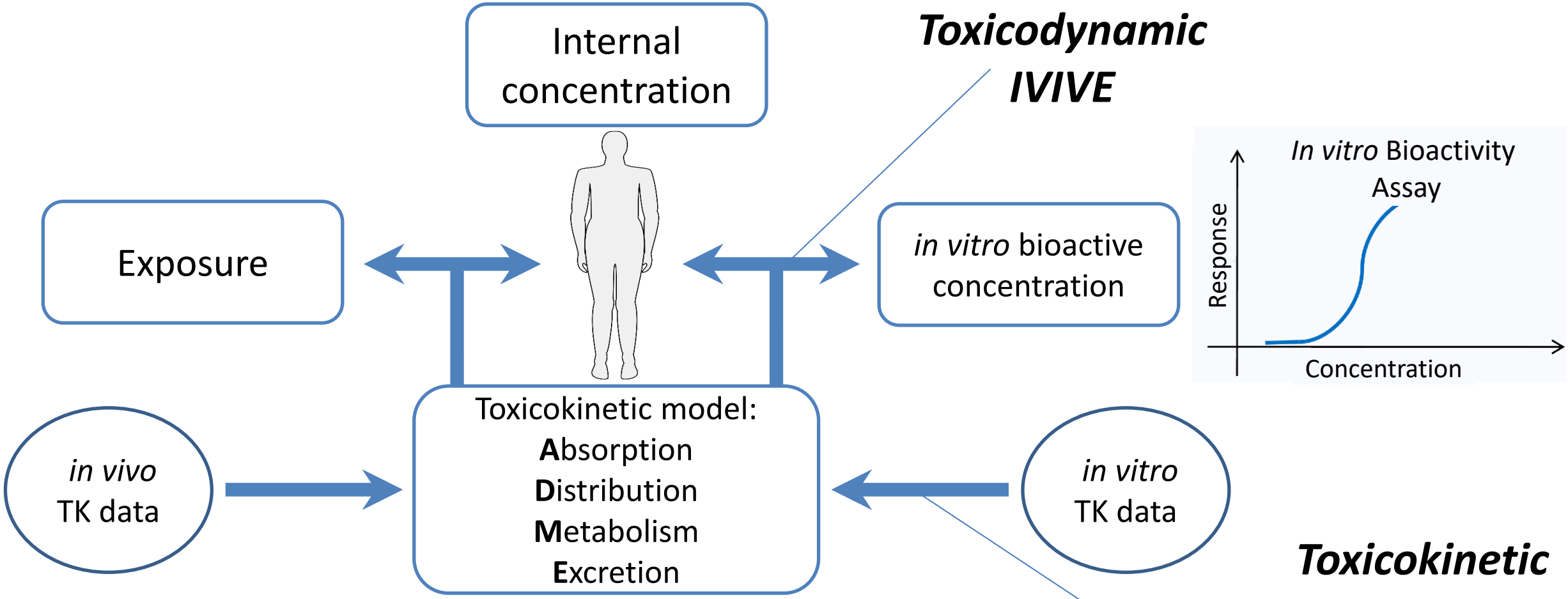




Bessems et al. (2014)

In Vitro-In Vivo Extrapolation (IVIVE)

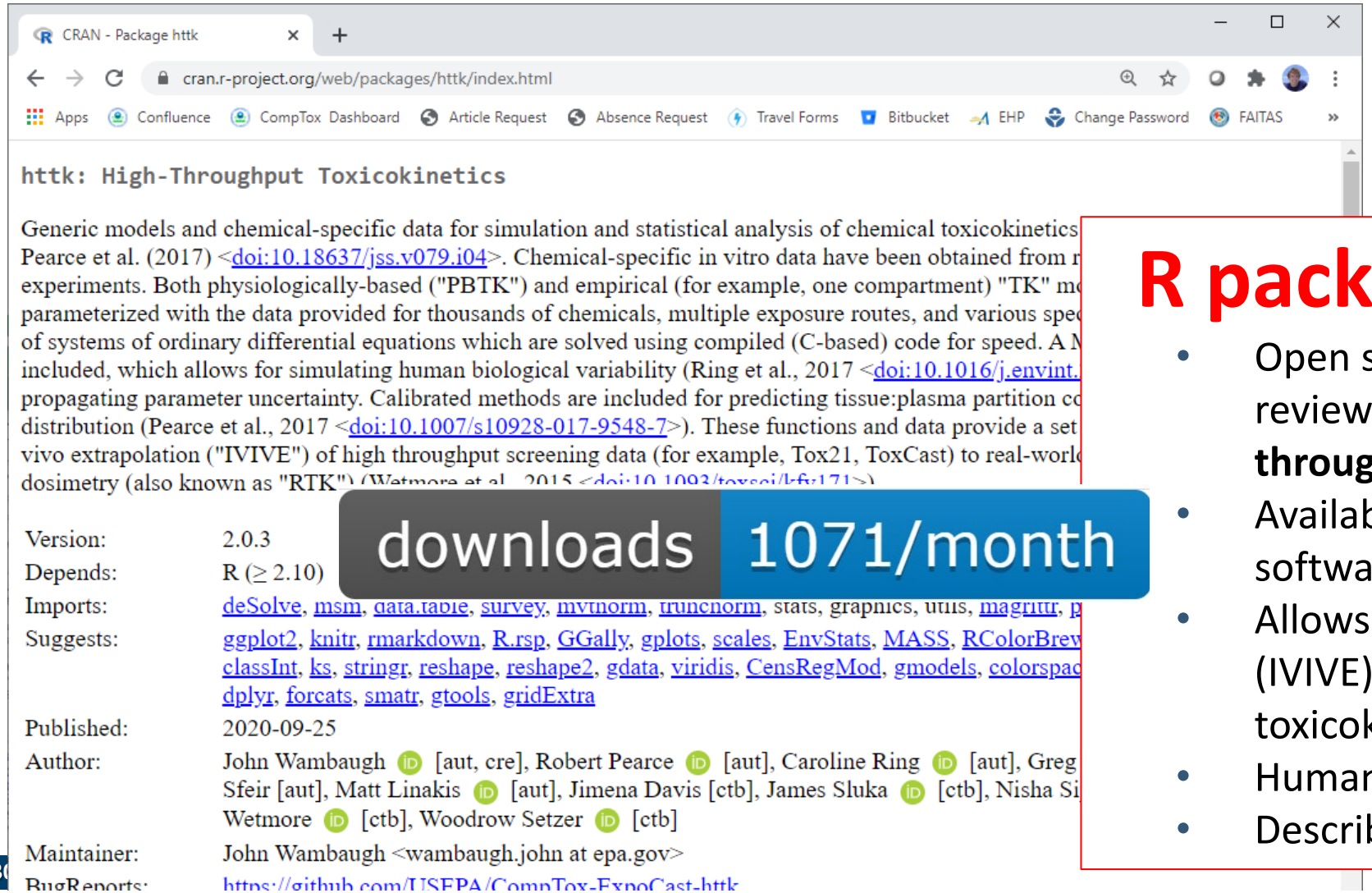
- Translation of *in vitro* high throughput screening requires chemical-specific toxicokinetic models
  - Needed for anywhere from dozens to thousands of chemicals



Breen et al. (submitted)

Open-Source Tools and Data for HTTK

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



R package "httk"

- Open source, transparent, and peer-reviewed tools and data for **high throughput toxicokinetics (httk)**
- Available publicly for free statistical software R
- Allows *in vitro-in vivo* extrapolation (IVIVE) and physiologically-based toxicokinetics (PBTK)
- Human-specific data for 987 chemicals
- Described in Pearce et al. (2017)

NAMs for Exposure Science

Exposure NAM Class	Description	Traditional Approach	NAM Makes Use of: Measurement	Toxicokinetics	Models	Descriptors	Evaluation	Machine Learning
Measurements	New techniques including screening analyses capable of detecting hundreds of chemicals present in a sample	Targeted (chemical-specific) analyses	-	●	●	●		●
Toxicokinetics	High throughput methods using *in vitro* data to generate chemical-specific models	Analyses based on *in vivo* animal studies	●	-		●		●
HTE Models	Models capable of making predictions for thousands of chemicals	Models requiring detailed, chemical- and scenario-specific information	●	●	-	●		
Chemical Descriptors	Informatic approaches for organizing chemical information in a machine-readable format	Tools targeted at single chemical analyses by humans				-		●
Evaluation	Statistical approaches that use the data from many chemicals to estimate the uncertainty in a prediction for a new chemical	Comparison of model predictions to data on a per chemical basis	●	●	●	●	-	●
Machine Learning	Computer algorithms to identify patterns	Manual Inspection of the Data	●	●		●		-
Prioritization	Integration of exposure and other NAMs to identify chemicals for follow-up study	Expert decision making	●	●	●	●	●	●

Wambaugh et al., (2019)

Chemical Property NAMs

United States Environmental Protection Agency


Contents lists available at ScienceDirect

Food and Chemical Toxicology

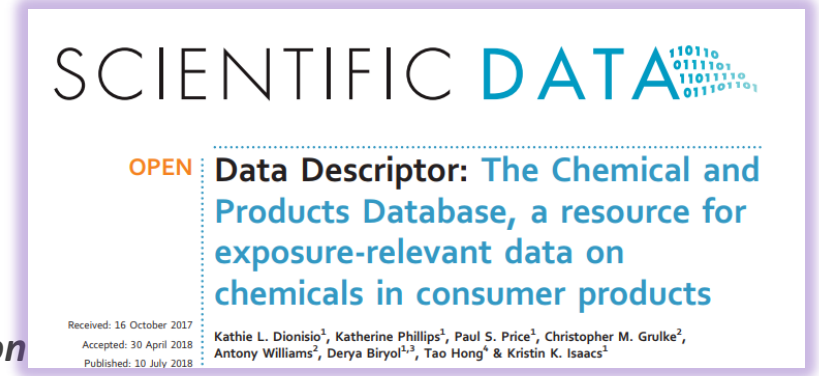
journal homepage: www.elsevier.com/locate/foodchemtox

Development of a consumer product ingredient database for chemical exposure screening and prioritization

M.-R. Goldsmith[a,*], C.M. Grulke[a], R.D. Brooks[b], T.R. Transue[c], Y.M. Tan[a], A. Frame[a,e], P.P. Egeghy[a], R. Edwards[d], D.T. Chang[a], R. Tornero-Velez[a], K. Isaacs[a], A. Wang[a,e], J. Johnson[a], K. Holm[a], M. Reich[f], J. Mitchell[g], D.A. Vallero[a], L. Phillips[a], M. Phillips[a], J.F. Wambaugh[a], R.S. Judson[a], T.J. Buckley[a], C.C. Dary[a]

MSDS Data

Occurrence and quantitative chemical composition


SCIENTIFIC DATA

OPEN

Data Descriptor: The Chemical and Products Database, a resource for exposure-relevant data on chemicals in consumer products

Received: 16 October 2017
Accepted: 30 April 2018
Published: 10 July 2018

Kathie L. Dionisio[1], Katherine Phillips[1], Paul S. Price[1], Christopher M. Grulke[2], Antony Williams[2], Derya Biryol[1,3], Tao Hong[4] & Kristin K. Isaacs[1]

Broad "index" of chemical uses


Contents lists available at ScienceDirect

Toxicology Reports

journal homepage: www.elsevier.com/locate/toxrep

Exploring consumer exposure pathways and patterns of use for chemicals in the environment

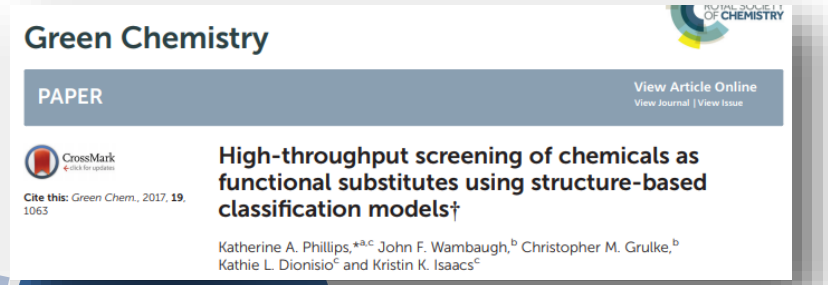
Kathie L. Dionisio[a], Alicia M. Frame[b,1], Michael-Rock Goldsmith[a,2], John F. Wambaugh[b], Alan Liddell[c,3], Tommy Cathey[d], Doris Smith[b], James Vail[b], Alexi S. Ernstoff[e], Peter Fantke[e], Olivier Jolliet[f],

CPCat

CPDat

EPA United States Environmental Protection Agency

Chemistry Dashboard


Green Chemistry

PAPER

View Article Online

View Journal | View Issue

CrossMark

Cite this: Green Chem., 2017, 19, 1063

High-throughput screening of chemicals as functional substitutes using structure-based classification models†

Katherine A. Phillips,*[a,c] John F. Wambaugh,[b] Christopher M. Grulke,[b] Kathie L. Dionisio[c] and Kristin K. Isaacs[c]

Functional Use Data

The roles that chemicals serve in products


Journal of Exposure Science and Environmental Epidemiology (2018) 28, 216–222

© 2018 Nature America, Inc., part of Springer Nature. All rights reserved 1559-0631/18

www.nature.com/jes

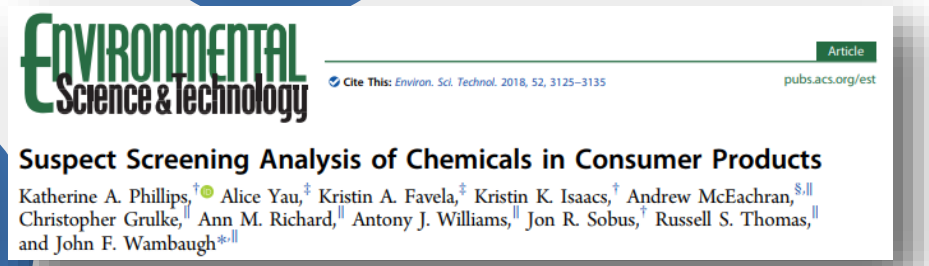
ORIGINAL ARTICLE

Consumer product chemical weight fractions from ingredient lists

Kristin K. Isaacs[1], Katherine A. Phillips[1], Derya Biryol[1,2], Kathie L. Dionisio[1] and Paul S. Price[1]

Ingredient Lists

Occurrence data


Environmental Science & Technology

Article

Cite This: Environ. Sci. Technol. 2018, 52, 3125−3135

pubs.acs.org/est

Suspect Screening Analysis of Chemicals in Consumer Products

Katherine A. Phillips,[†] Alice Yau,[‡] Kristin A. Favela,[‡] Kristin K. Isaacs,[†] Andrew McEachran,[§,∥] Christopher Grulke,[∥] Ann M. Richard,[∥] Antony J. Williams,[∥] Jon R. Sobus,[†] Russell S. Thomas,[∥] and John F. Wambaugh*[,∥]

Measured Data

Measurement of chemicals in consumer products

https://comptox.epa.gov/dashboard

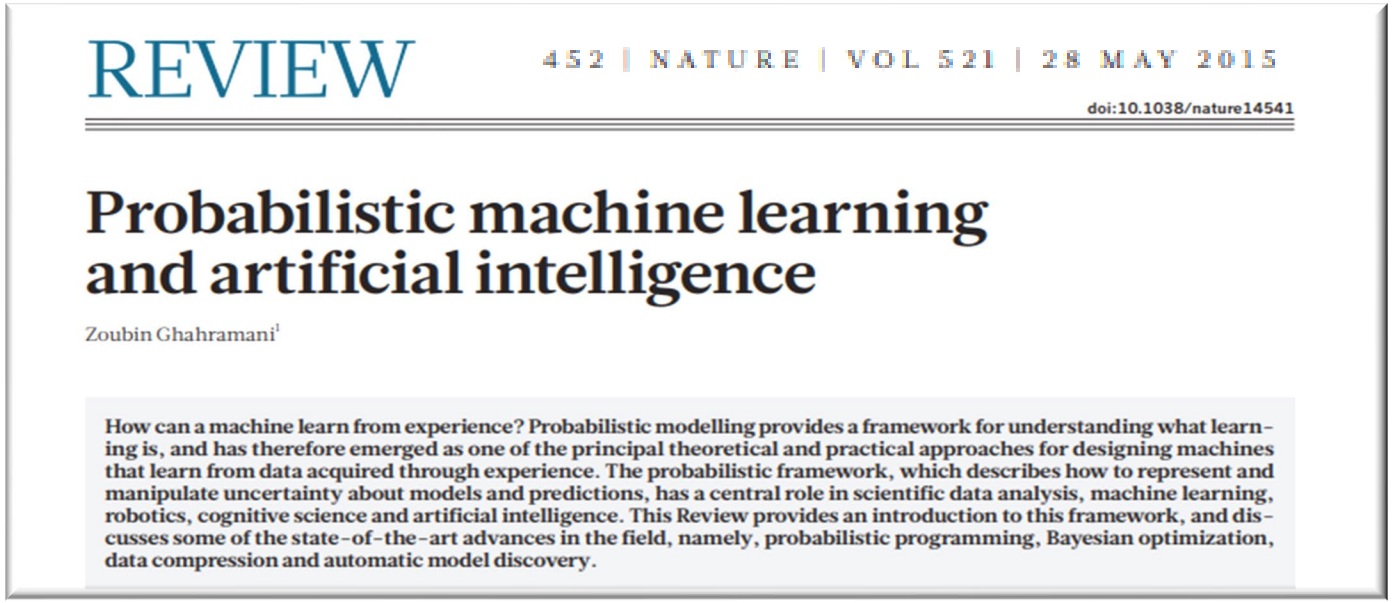
NAMs for Exposure Science

Exposure NAM Class	Description	Traditional Approach	NAM Makes Use of					
			Measurement	Toxicokinetics	Models	Descriptors	Evaluation	Machine Learning
Measurements	New techniques including screening analyses capable of detecting hundreds of chemicals present in a sample	Targeted (chemical-specific) analyses	-	●	●	●		●
Toxicokinetics	High throughput methods using *in vitro* data to generate chemical-specific models	Analyses based on *in vivo* animal studies	●	-		●		●
HTE Models	Models capable of making predictions for thousands of chemicals	Models requiring detailed, chemical- and scenario-specific information	●	●	-	●		
Chemical Descriptors	Informatic approaches for organizing chemical information in a machine-readable format	Tools targeted at single chemical analyses by humans				-		●
Evaluation	Statistical approaches that use the data from many chemicals to estimate the uncertainty in a prediction for a new chemical	Comparison of model predictions to data on a per chemical basis	●	●	●	●	-	●
Machine Learning	Computer algorithms to identify patterns	Manual Inspection of the Data	●	●		●		-
Prioritization	Integration of exposure and other NAMs to identify chemicals for follow-up study	Expert decision making	●	●	●	●	●	●

Wambaugh et al., (2019)

Machine Learning: A Subset of Artificial Intelligence

"…machine learning can be thought of as inferring plausible models to explain observed data."



REVIEW

452 | NATURE | VOL 521 | 28 MAY 2015

doi:10.1038/nature14541

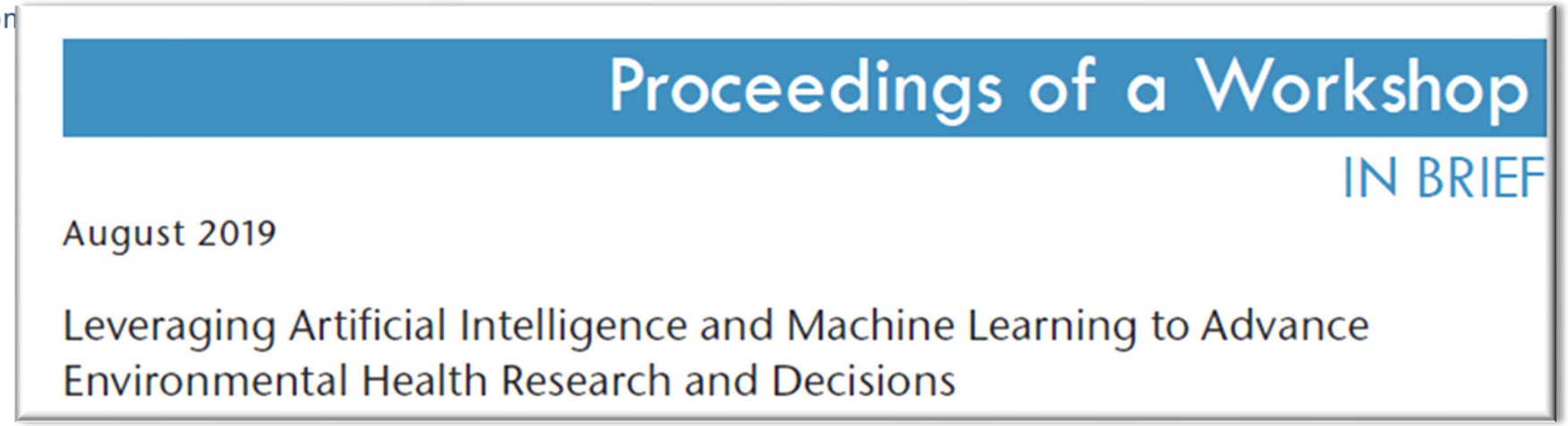
Probabilistic machine learning and artificial intelligence

Zoubin Ghahramani[1]

How can a machine learn from experience? Probabilistic modelling provides a framework for understanding what learning is, and has therefore emerged as one of the principal theoretical and practical approaches for designing machines that learn from data acquired through experience. The probabilistic framework, which describes how to represent and manipulate uncertainty about models and predictions, has a central role in scientific data analysis, machine learning, robotics, cognitive science and artificial intelligence. This Review provides an introduction to this framework, and discusses some of the state-of-the-art advances in the field, namely, probabilistic programming, Bayesian optimization, data compression and automatic model discovery.

At the EPA we are applying publicly available machine learning algorithms to bridge data gaps and draw inferences from complex data sets.

Machine Learning in Environmental Decision-Making



- ***National Academies Workshop, June 2019***
  - *"Machine learning algorithms can analyze large volumes of complex data to find patterns and make predictions, often exceeding the accuracy and efficiency of people who are attempting the same task."*
  - Highlighted areas of environmental health for which AI and machine learning could help, including:
    - Predicting the toxicology of chemicals
    - Characterizing the exposome

Machine Learning NAMS

United States Environmental Protection Agency

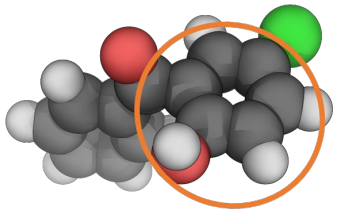
Chemical Functional Use Database (FUSE)

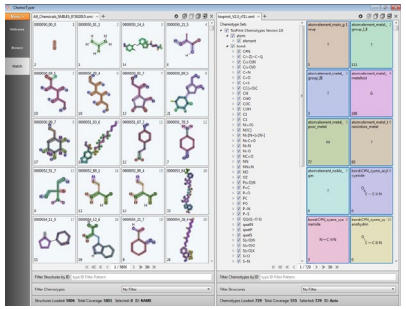
Positive Examples

Negative Examples

Random Forest Classification Models (Breiman, 2001) with five-fold cross validation

Chemical Structure and Property Descriptors





Successful Model

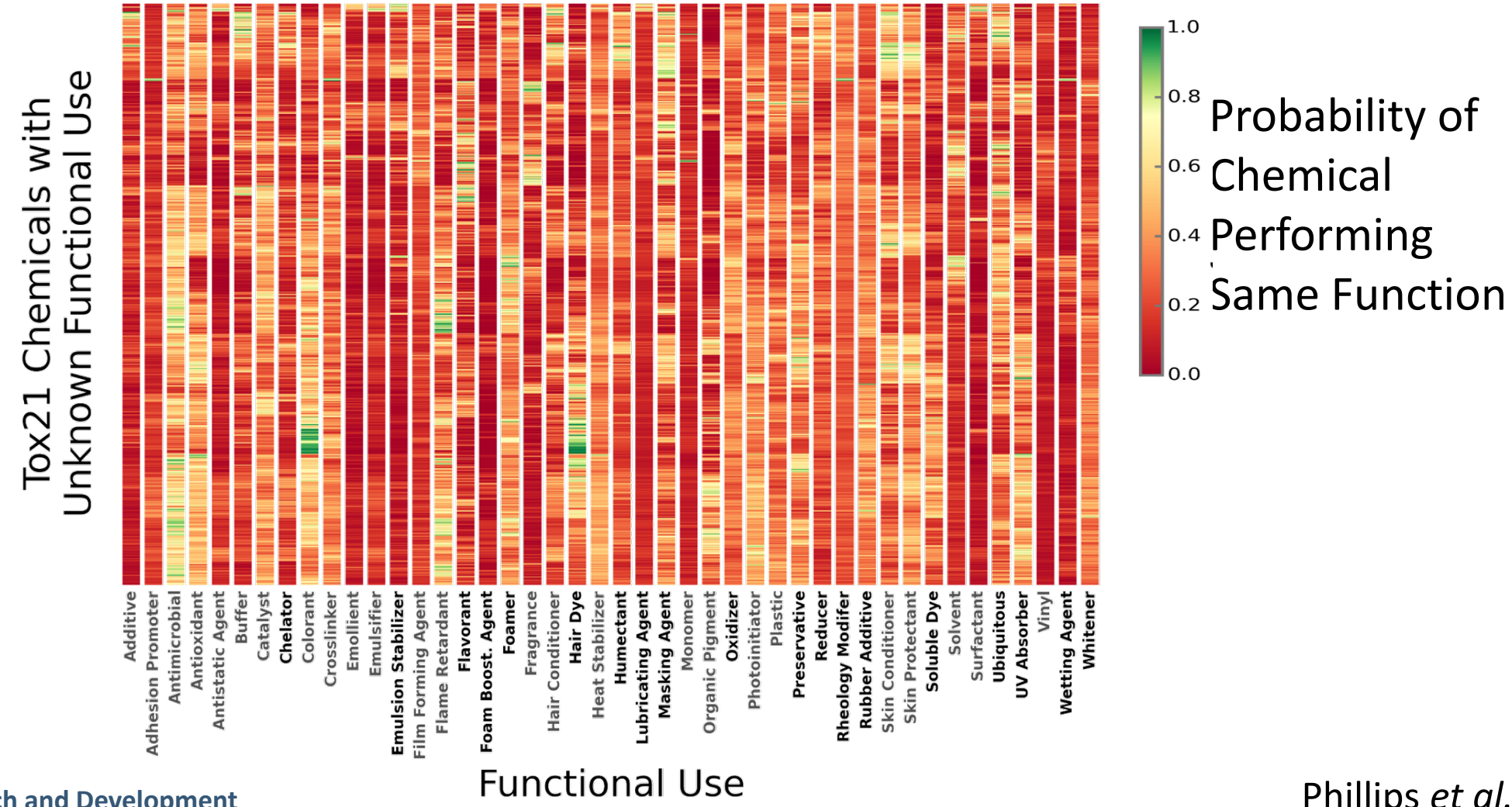
Failed Model

Probabilistic Predictions of Potential Chemical Uses

Phillips *et al.* (2017)

Screening for Alternatives By Function and Bioactivity

Combine high throughput screening data and chemical use prediction:



Phillips *et al.* (2017)

NAMs for Exposure Science

Exposure NAM Class	Description	Traditional Approach	NAM Makes Use of: Measurement	Toxicokinetics	Models	Descriptors	Evaluation	Machine Learning
Measurements	New techniques including screening analyses capable of detecting hundreds of chemicals present in a sample	Targeted (chemical-specific) analyses	-	●	●	●		●
Toxicokinetics	High throughput methods using *in vitro* data to generate chemical-specific models	Analyses based on *in vivo* animal studies	●	-		●		●
HTE Models	Models capable of making predictions for thousands of chemicals	Models requiring detailed, chemical- and scenario-specific information	●	●	-	●		
Chemical Descriptors	Informatic approaches for organizing chemical information in a machine-readable format	Tools targeted at single chemical analyses by humans				-		●
Evaluation	Statistical approaches that use the data from many chemicals to estimate the uncertainty in a prediction for a new chemical	Comparison of model predictions to data on a per chemical basis	●	●	●	●	-	●
Machine Learning	Computer algorithms to identify patterns	Manual Inspection of the Data	●	●		●		-
Prioritization	Integration of exposure and other NAMs to identify chemicals for follow-up study	Expert decision making	●	●	●	●	●	●

Wambaugh et al., (2019)

Published and Ongoing Non-Targeted Analysis (NTA) Studies in the ExpoCast Project

United States Environmental Protection Agency

Source and Release

Pilot: 20 Consumer Product Categories

Phillips *et al.*, *Env. Sci. Tech*. 2018

Recycled Consumer Materials

Lowe *et al.*, *Submitted*

Consumer Product Emissions from Different Substrates

Fate and Transport

Residential Air

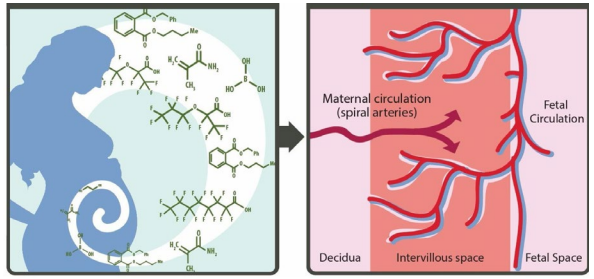
Residential Dust

Rager *et al.*, *Env. Int.*, 2016

Exposure

Pooled Human Blood

Human Placenta



Rager *et al.*, *Repro. Tox.* , 2020

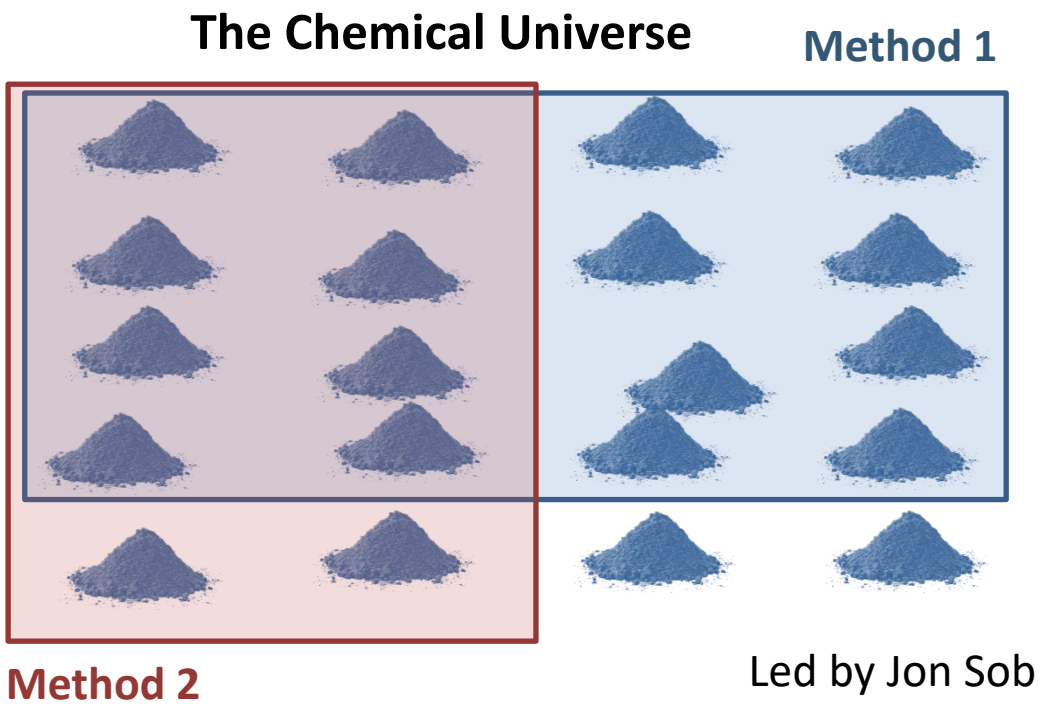
Emerging Science: How can we **quantify** concentrations of chemicals in media using NTA?

Slide from Kristin Isaacs

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,
Seth Newton and Elin Ulrich

- Phase 1:
  - Collaborators provided 10 mixtures of 100-400 ToxCast chemicals each
  - Mass spectrometry equipment vendors provided with individual chemical standards

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

Ulrich et al. (2019)

NAMs for Exposure Science

Exposure NAM Class	Description	Traditional Approach	NAM Makes Use of: Measurement	Toxicokinetics	Models	Descriptors	Evaluation	Machine Learning
Measurements	New techniques including screening analyses capable of detecting hundreds of chemicals present in a sample	Targeted (chemical-specific) analyses	-	●	●	●		●
Toxicokinetics	High throughput methods using *in vitro* data to generate chemical-specific models	Analyses based on *in vivo* animal studies	●	-		●		●
HTE Models	Models capable of making predictions for thousands of chemicals	Models requiring detailed, chemical- and scenario-specific information	●	●	-	●		
Chemical Descriptors	Informatic approaches for organizing chemical information in a machine-readable format	Tools targeted at single chemical analyses by humans				-		●
Evaluation	Statistical approaches that use the data from many chemicals to estimate the uncertainty in a prediction for a new chemical	Comparison of model predictions to data on a per chemical basis	●	●	●	●	-	●
Machine Learning	Computer algorithms to identify patterns	Manual Inspection of the Data	●	●		●		-
Prioritization	Integration of exposure and other NAMs to identify chemicals for follow-up study	Expert decision making	●	●	●	●	●	●

Wambaugh et al., (2019)

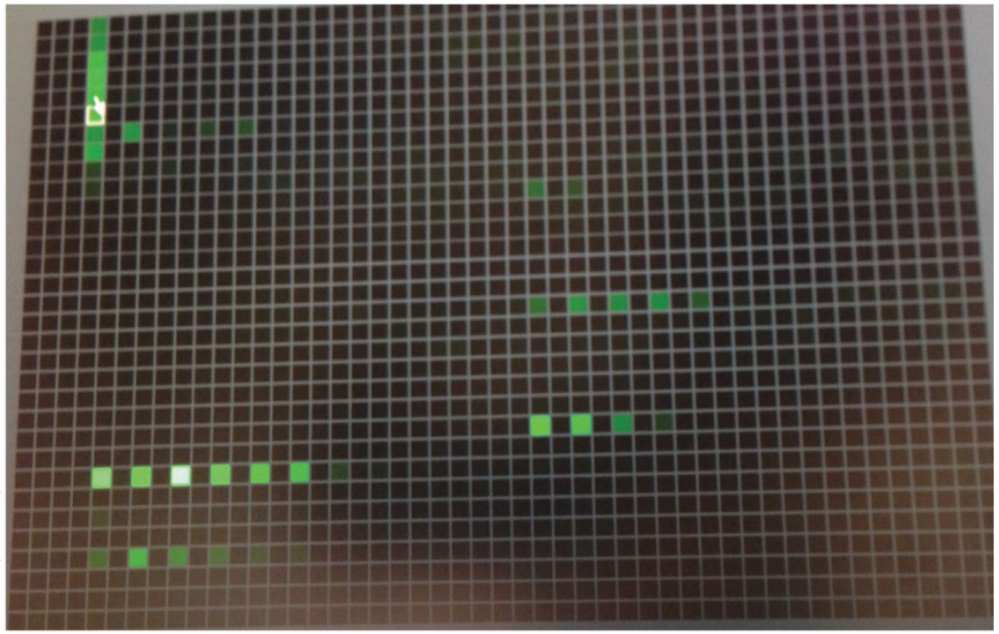
What is "High Throughput"?

- Tox21: Testing one assay across 10,000 chemicals takes 1-2 days, but only 50 assays have been developed so far that can run that fast

- ToxCast: ~1100 off-the-shelf (pharma) assay-endpoints tested for up to 4,000 chemicals over the past decade, now developing new assays as well

HTS tox assays often use single readout, such as fluorescence, across many chemicals, measuring concentration for toxicokinetics or exposure requires chemical-specific methods…

Kaewkhaw et al. (2016)



What is "High Throughput"?

- Tox21: Testing one assay across 10,000 chemicals takes 1-2 days, but only 50 assays have been developed so far that can run that fast

- ToxCast: ~1100 off-the-shelf (pharma) assay-endpoints tested for up to 4,000 chemicals over the past decade, now developing new assays as well

HTS tox assays often use single readout, such as fluorescence, across many chemicals, measuring concentration for toxicokinetics or exposure requires chemical-specific methods…

- ExpoCast: Ring et al. made *in silico* predictions for ~480,000 chemicals from structure, but based on NHANES monitoring for ~120 chemicals
  - Quantitative non-targeted analysis (NTA) may eventually provide greater evaluation data to reduce uncertainty

- HTTK: *In vitro* data on 944 chemicals collected for humans, starting with Rotroff et al. (2010)
  - Work continues to develop *in silico* tools, for example Sipes et al. (2016)

Our work is not done…

How Do We Know if a Model is a "High Throughput Exposure (HTE) Model"?

To be considered an HTE model, a model must:

1. Be applicable to and capable of **handling many chemicals** with minimal descriptive information
2. Cover one or more relevant **exposure routes** (for example, inhalation, food ingestion, mouthing, and dermal contact) and **sources** (for example, industrial and residential use), accounting for the influential parameters relevant for the considered pathways
3. **Allow for integration** with models for other pathways
4. Be **scientifically plausible**, respecting mass-balance principles and accounting for competing processes (for example, volatilization versus dermal uptake)
5. Allow for the assessment of **interindividual and intraindividual variation** in exposure and impact of such variation on acute and chronic doses as the required input data become available
6. Be amenable to integration within statistical frameworks that **quantify uncertainty** for propagation into risk evaluations
7. **Remain parsimonious**, that is, no more complicated than necessary to reflect the data

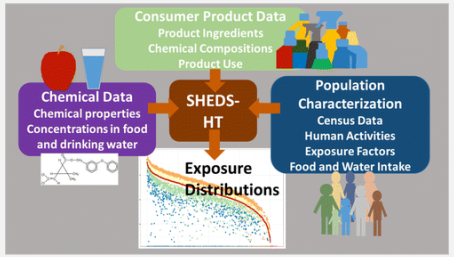
Adapted by Wambaugh et al. (2019) from Huang and Jolliet (2016)

High Throughput Exposure (HTE) Models for Key Pathways

United States Environmental Protection Agency

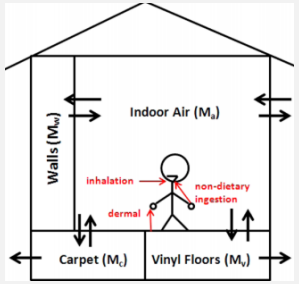
Consumer (Near-Field) Pathways

SHEDS-HT *(Isaacs et al., 2014)*



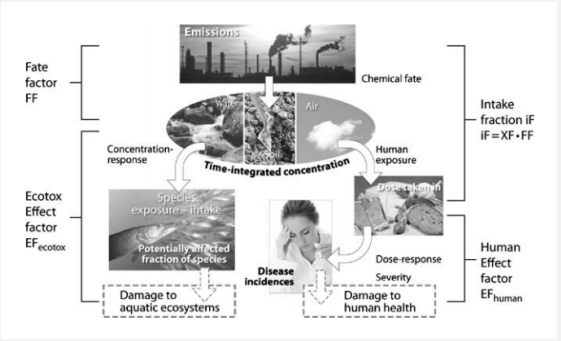
RAIDAR-ICE *(Li et al., 2018)*

FINE *(Shin et al., 2015)*

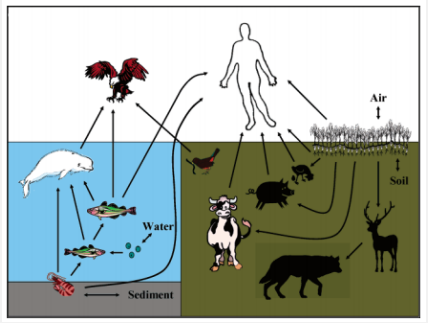


Ambient (Far-Field) Pathways

UseTox *(Rosenbaum et al., 2008)*

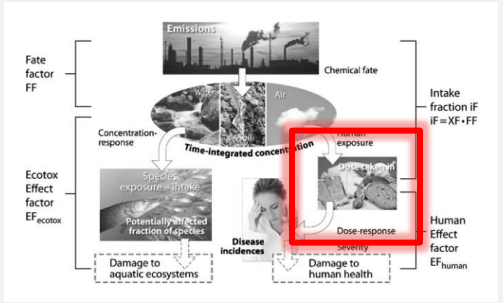


RAIDAR *(Arnot et al., 2006, 2008)*

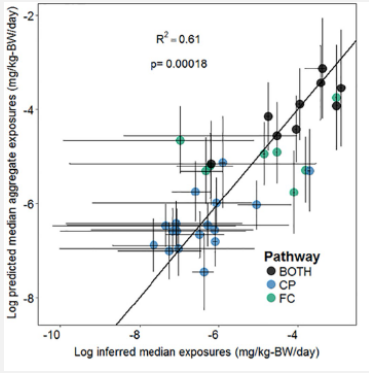


Dietary Pathways

UseTox *(Rosenbaum et al. (2008)*



SHEDS-HT *(Biryol et al., 2017)*



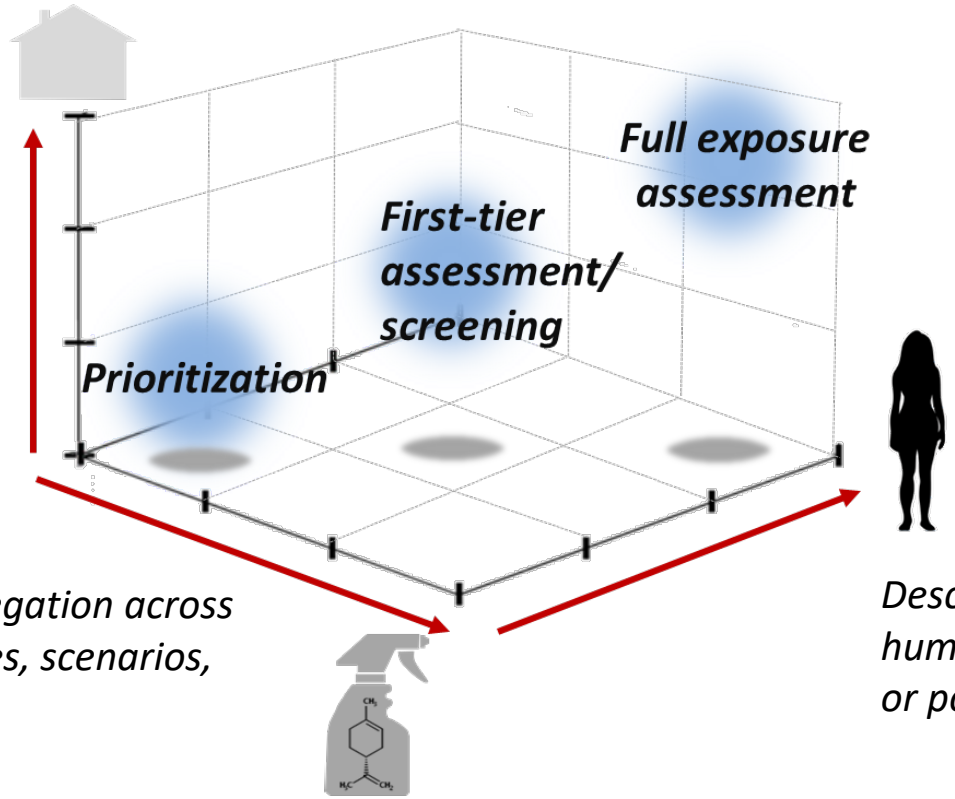
Slide from Kristin Isaacs

Fit-for-Purpose Exposure Modeling Frameworks

United States Environmental Protection Agency

Increasing Complexity

Mechanistic description of the built environment and exposure processes, including temporal variability



Level of aggregation across sources, routes, scenarios, chemicals

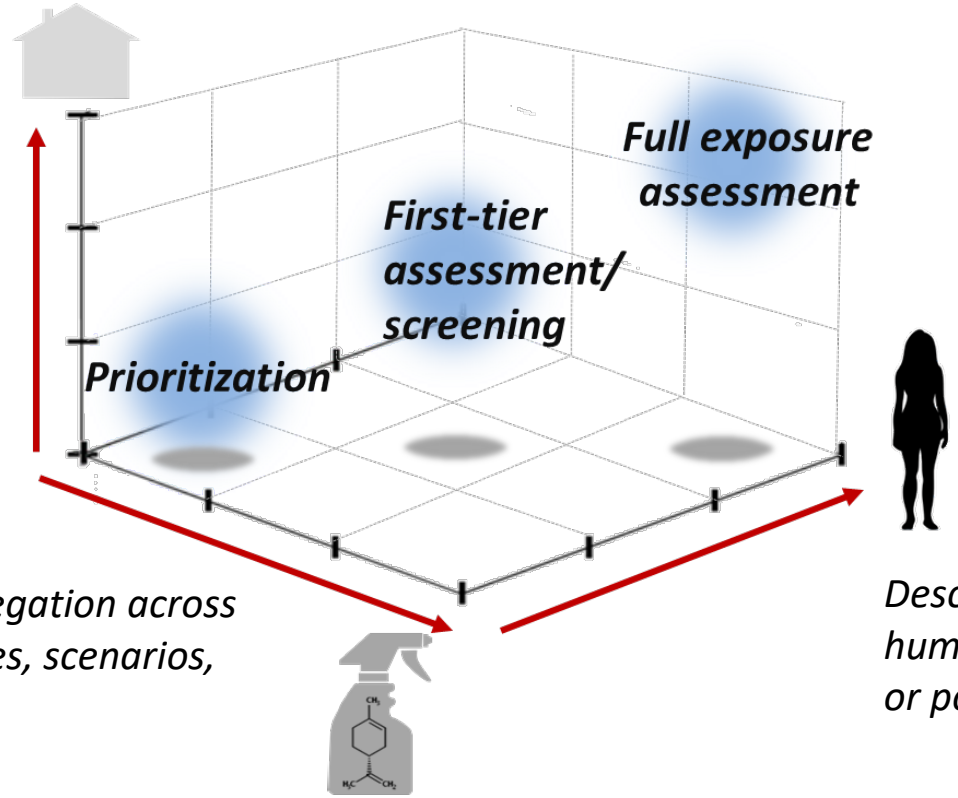
Description of human behavior or population

Slide from Kristin Isaacs

Fit-for-Purpose Exposure Modeling Frameworks

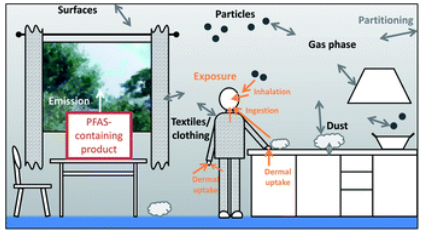
Increasing Complexity

Mechanistic description of the built environment and exposure processes, including temporal variability

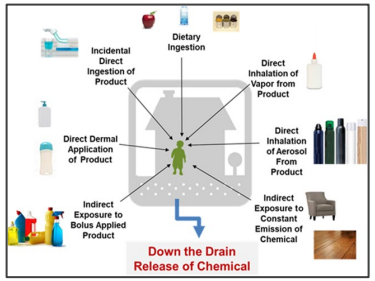


Level of aggregation across sources, routes, scenarios, chemicals

Description of human behavior or population

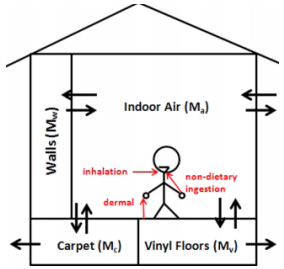


Eichler and Little, 2020



SHEDS-HT, Isaacs et al., 2014

Li et al., 2018



FINE, Shin et al., 2015

- Models of different levels of complexity have **overlapping data needs**
- They also share some **universal challenges**

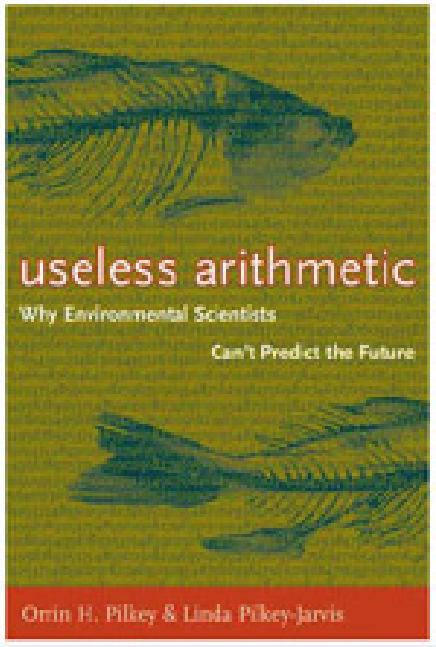
Consumer Exposure Model

EPA, 2019

Slide from Kristin Isaacs

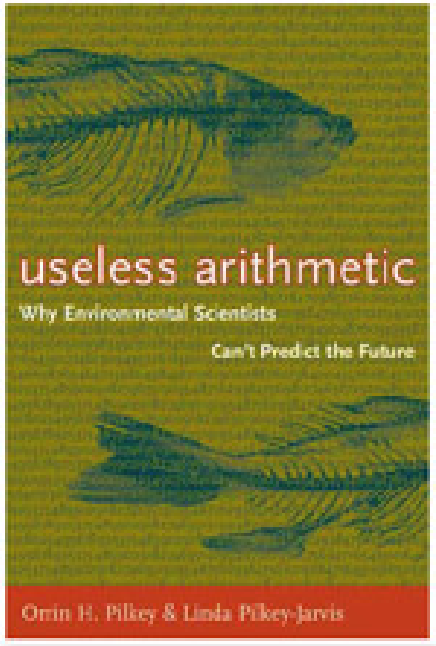
How to Make Good Forecasts

Adapted from Nate Silver



Orrin Pilkey &
Olinda Pilkey-Jarvis (2007)

How to Make Good Forecasts
Adapted from Nate Silver

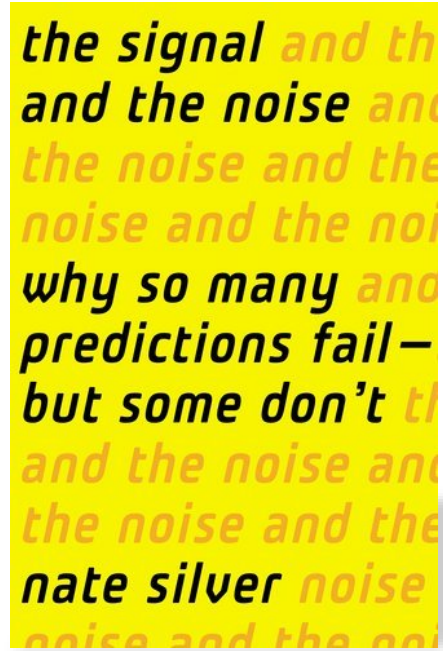


Orrin Pilkey &
Olinda Pilkey-Jarvis (2007)

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

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

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

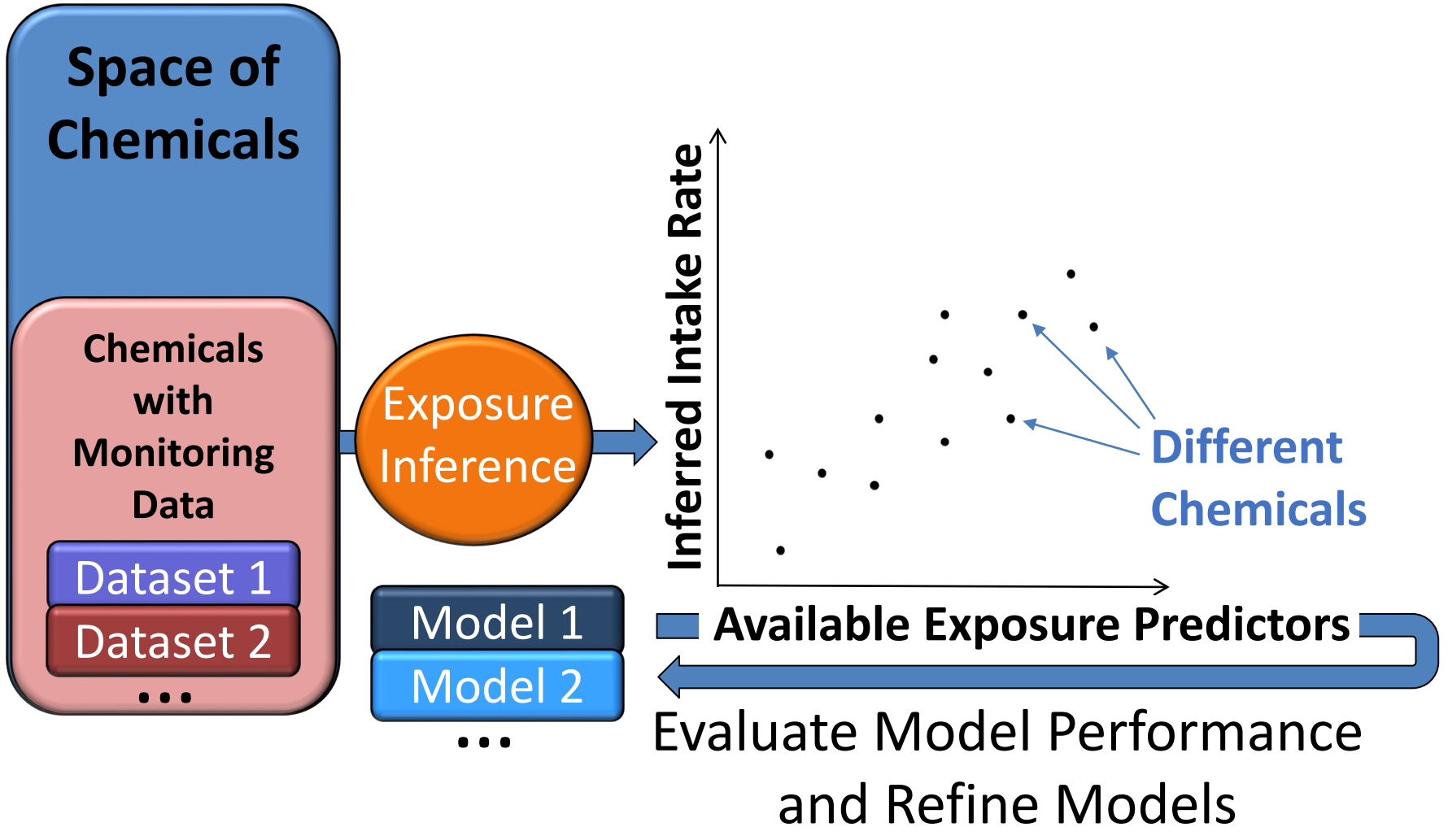


Nate Silver (2012)

In Nate Silver's terminology:
a ***prediction*** is a specific statement
a ***forecast*** is a probabilistic statement

Evaluating Exposure Models with the SEEM Framework

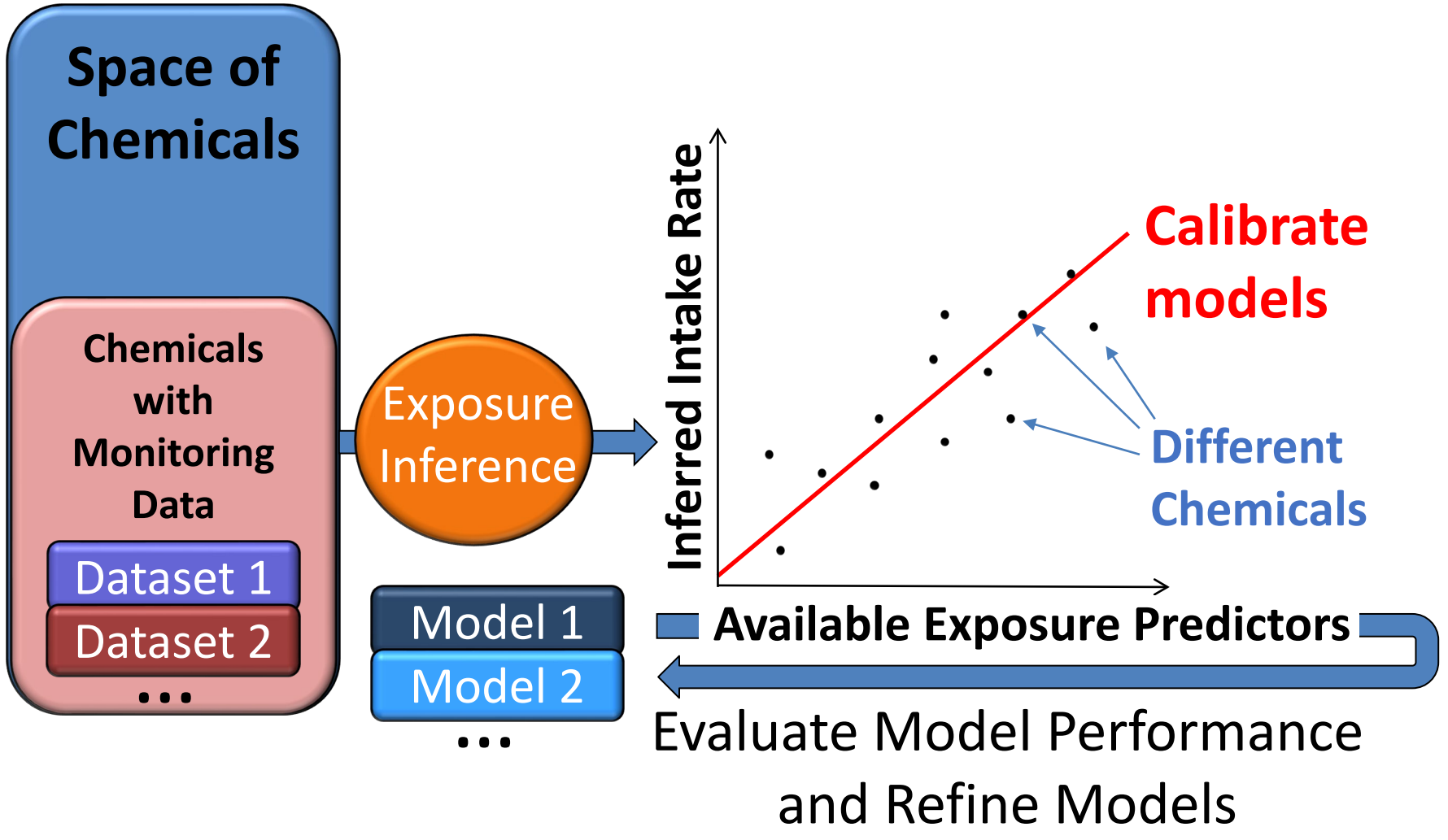
- We use Bayesian methods to incorporate multiple models into consensus predictions for 1000s of chemicals within the **Systematic Empirical Evaluation of Models (SEEM)** (Wambaugh et al., 2013, 2014; Ring et al., 2018)



Wambaugh et al., 2018

Evaluating Exposure Models with the SEEM Framework

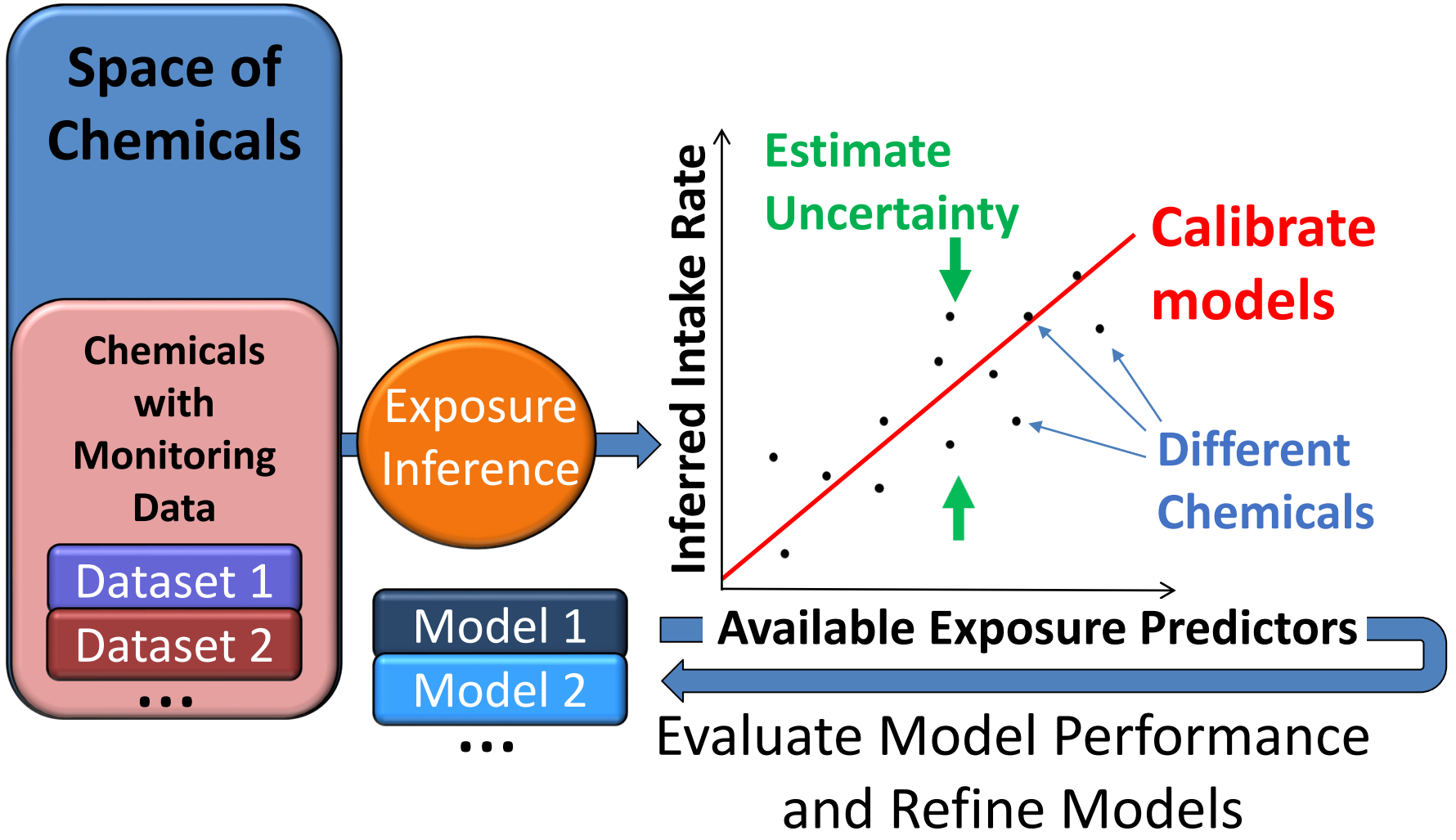
- We use Bayesian methods to incorporate multiple models into consensus predictions for 1000s of chemicals within the **Systematic Empirical Evaluation of Models (SEEM)** (Wambaugh et al., 2013, 2014; Ring et al., 2018)



Wambaugh et al., 2018

Evaluating Exposure Models with the SEEM Framework

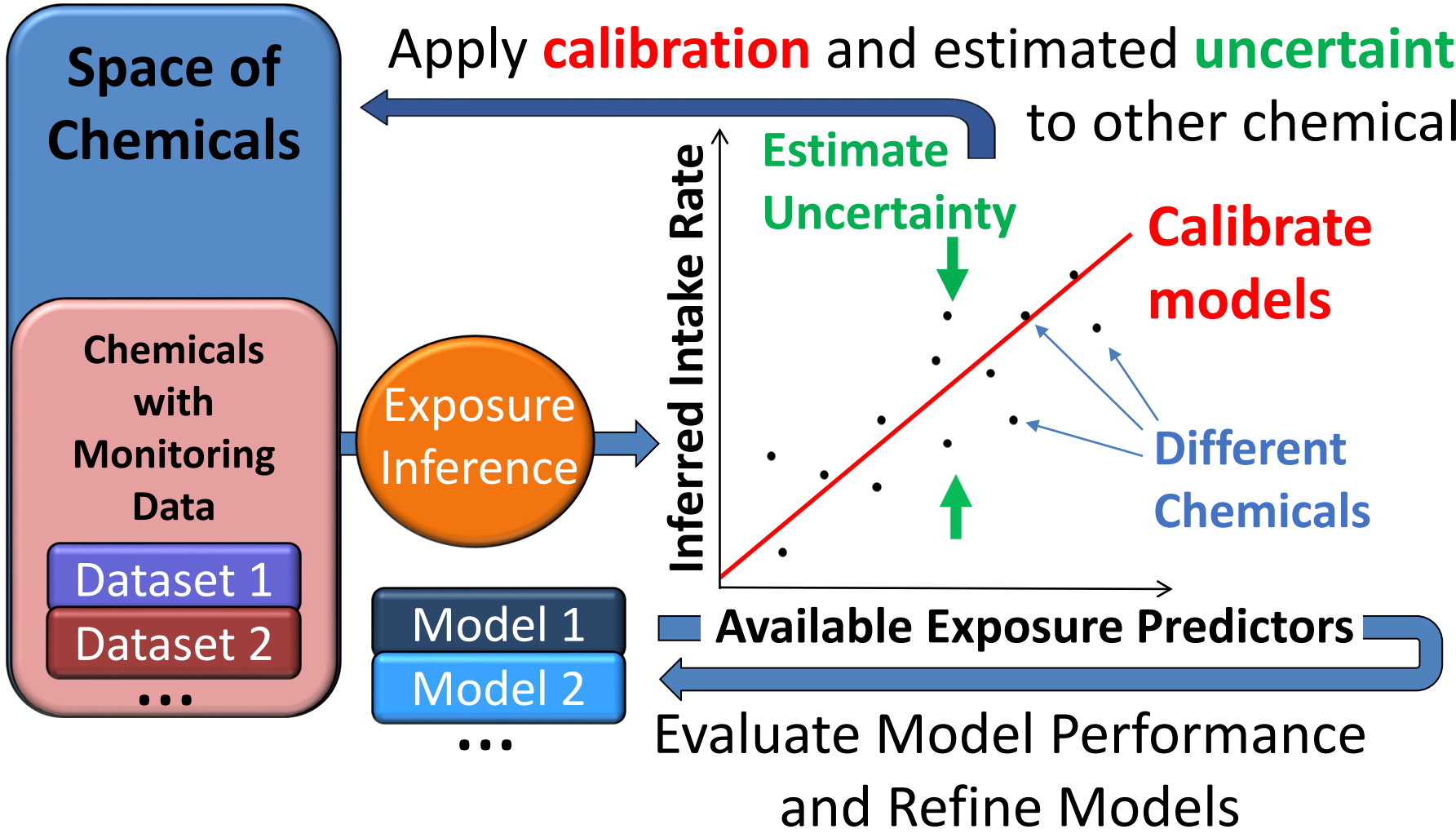
- We use Bayesian methods to incorporate multiple models into consensus predictions for 1000s of chemicals within the **Systematic Empirical Evaluation of Models (SEEM)** (Wambaugh et al., 2013, 2014; Ring et al., 2018)



Wambaugh et al., 2018

Evaluating Exposure Models with the SEEM Framework

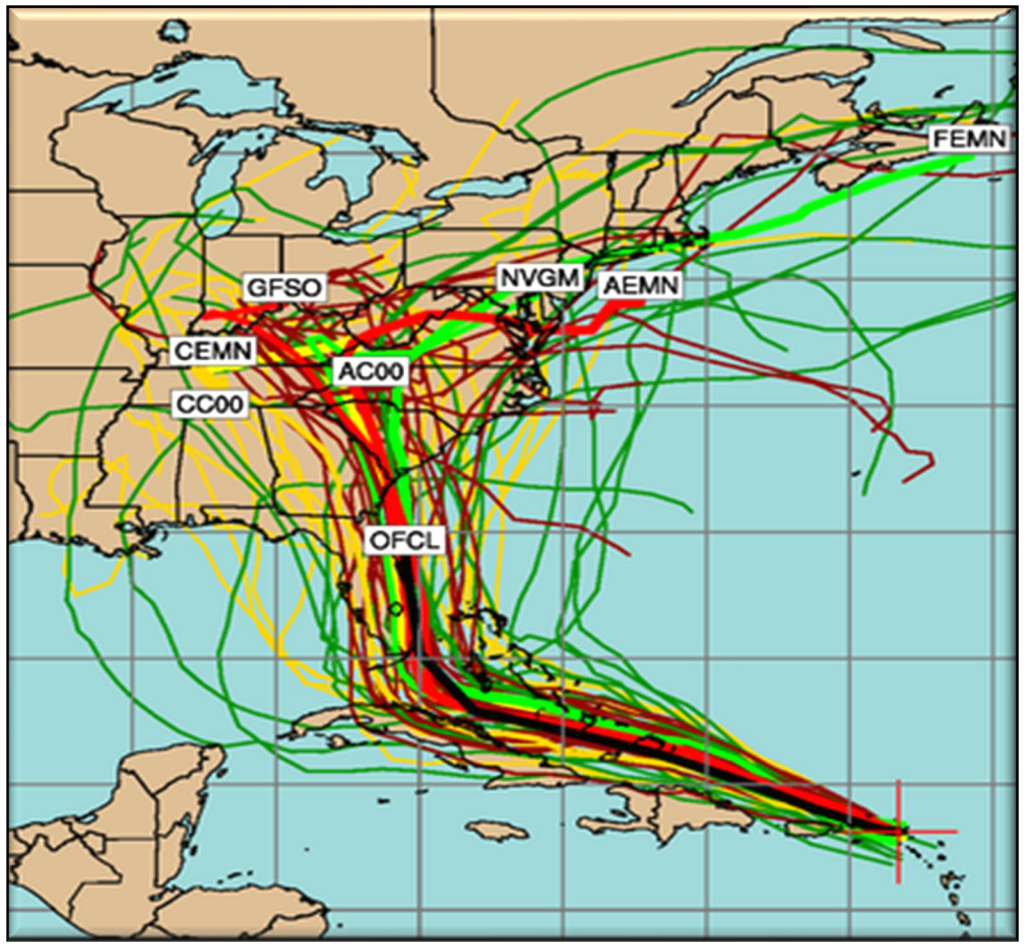
- We use Bayesian methods to incorporate multiple models into consensus predictions for 1000s of chemicals within the **Systematic Empirical Evaluation of Models (SEEM)** (Wambaugh et al., 2013, 2014; Ring et al., 2018)



Wambaugh et al., 2018

Ensemble Predictions

- We can use ensemble methods to make more stable models and characterize uncertainty
- "Ensemble methods are learning algorithms that construct a set of classifiers and then classify new data points by taking a (weighted) vote of their predictions." Dietterich (2000)
- Ensemble systems have proven themselves to be very effective and extremely versatile in a broad spectrum of problem domains and real-world applications (Polikar, 2012)
- Ensemble learning techniques in the machine learning paradigm can be used to integrate predictions from multiple tools. (Pradeep, 2016)



Hurricane Path Prediction is an Example of Integrating Multiple Models

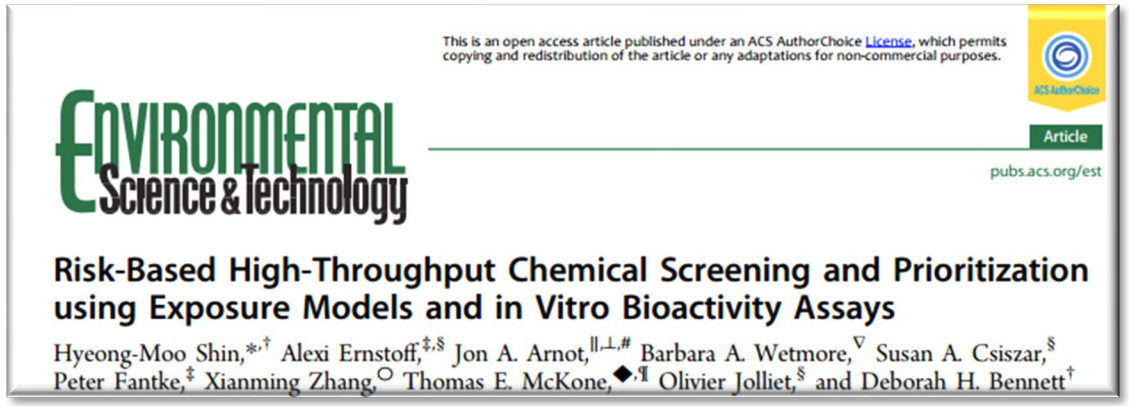
Knowledge of Exposure Pathways Limits High Throughput Exposure Models

- Wambaugh et al. (2014) found that “pesticide inerts” had higher than average levels in biomonitoring data, while “pesticide actives” had lower than average

- Pesticide inerts have many other uses, but there are more stringent reporting requirements for pesticides
  - **Exposure is occuring by other pathways**

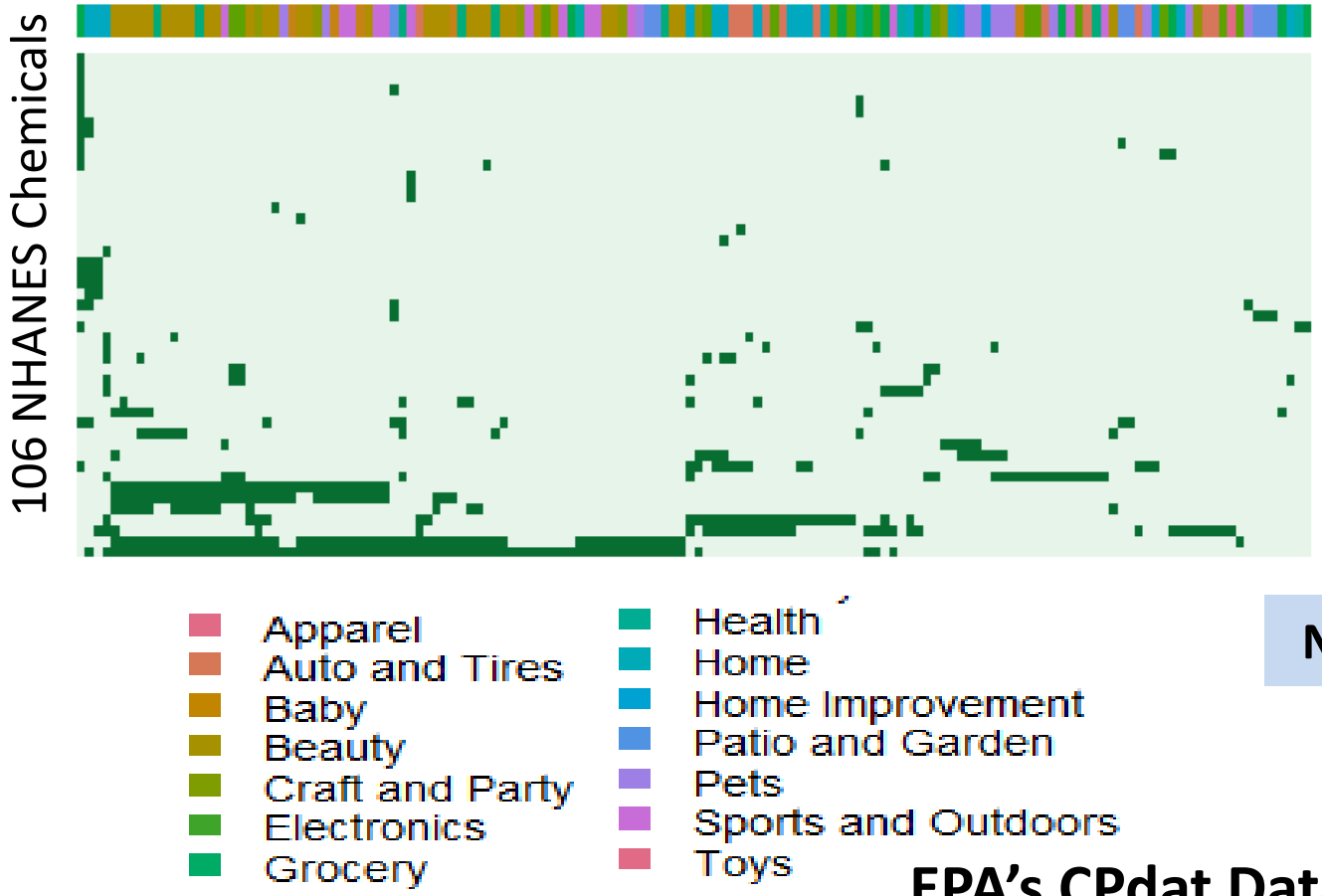
- But we don’t always know how chemicals are used:

  “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.”

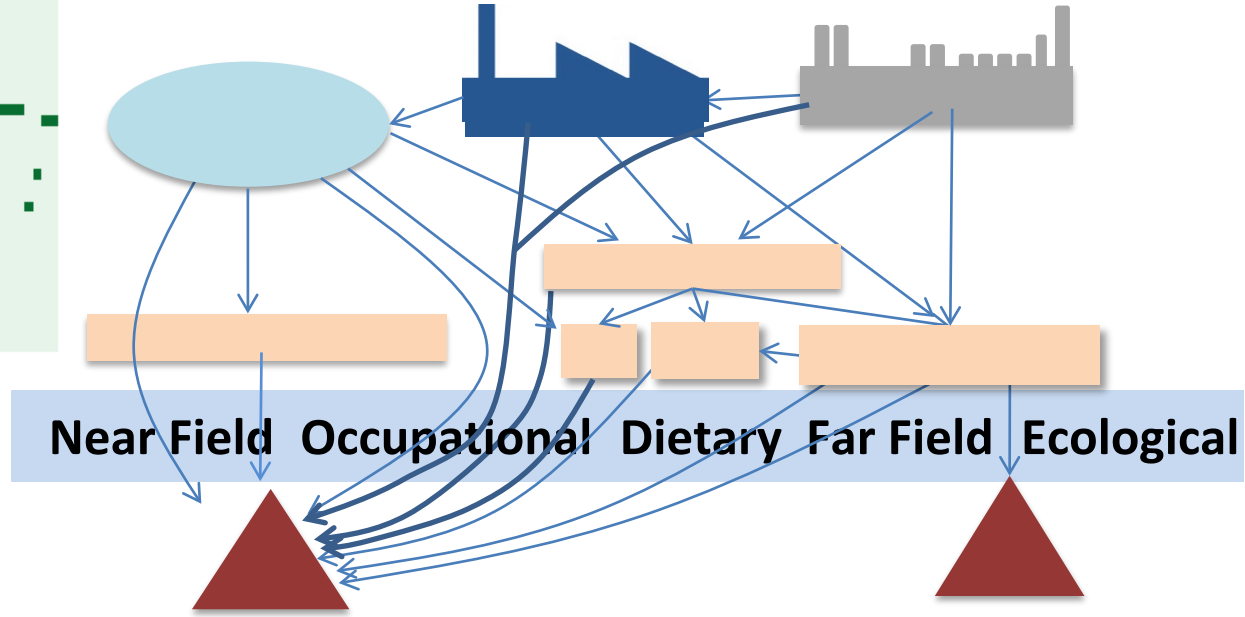


Chemical Use Identifies Relevant Pathways

>2000 chemicals with Material Safety Data Sheets (MSDS) in CPCPdb (Goldsmith *et al.*, 2014)



Some pathways have much higher average exposures!



EPA's CPdat Database provides chemical-product information:

https://comptox.epa.gov/dashboard/

Machine Learning to 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.* (2018)

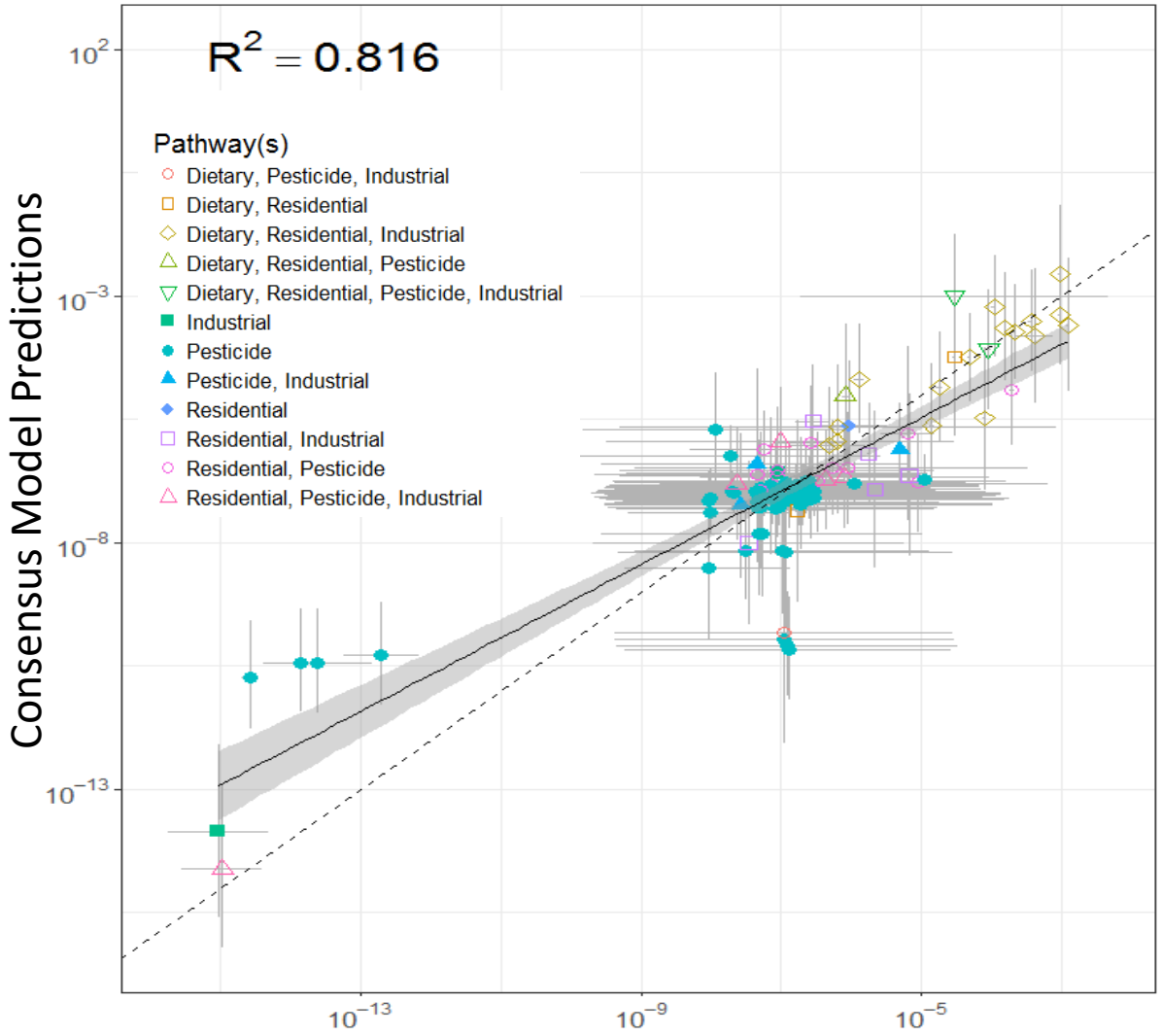
SEEM3 Collaboration

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	Pathway(s)
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

Ring et al. (2018)

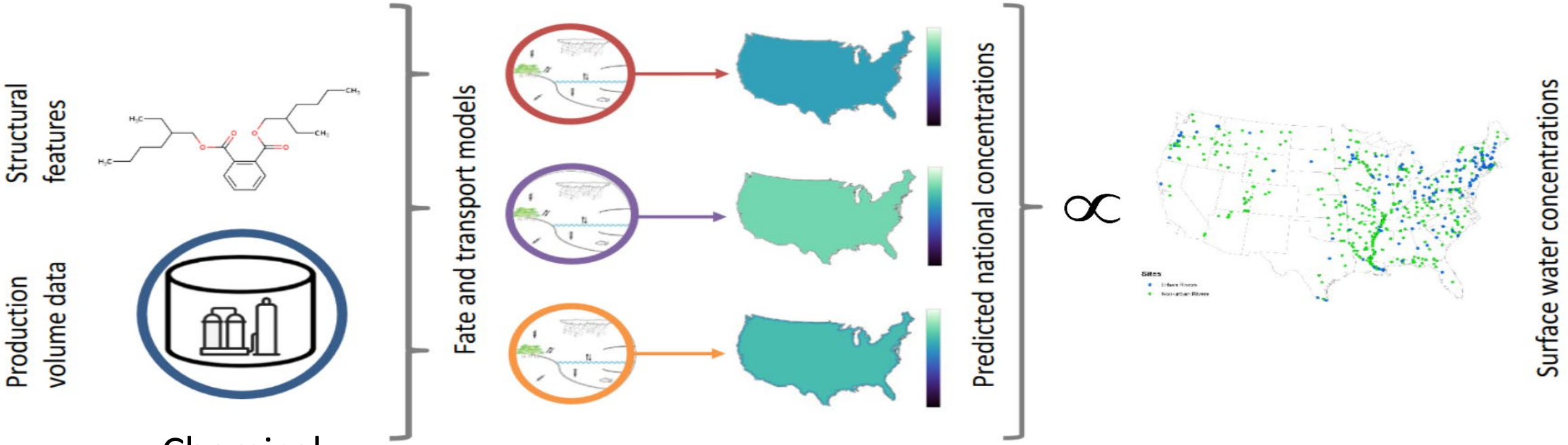
SEEM3: Pathway-Based Consensus Modeling



- SEEM3 consensus model provides estimates of human median intake rate (mg/kg/day) for nearly 500,000 chemicals via the CompTox Chemicals Dashboard (http://comptox.epa.gov/dashboard)
- SEEM3 first predicts relevant exposure pathways from chemical structure – model predictions are then weighted according to the models' abilities to explain NHANES data
- We rely on pathway determinations from CPDat
- We rely on NHANES biomonitoring data
  - 2014 FIFRA Scientific Advisory Panel identified need for broader sets of evaluation data

Ring et al., 2018

EcoSEEM Metamodel for Surface Water Chemical Concentrations



Chemical-specific information (*l*) → Release (loading) and fate predictors (*p*) → **EcoSEEM metamodel** (*m*) ← USGS/EPA water monitoring data (*y*)

Sayre et al,
in preparation

$$\ln y_i = m_0 + \sum_{j=1}^{n_j} \sum_{k=1}^{n_{kj}} m_{jk} \ln\left(l_{ji} p_{ki}\right)$$

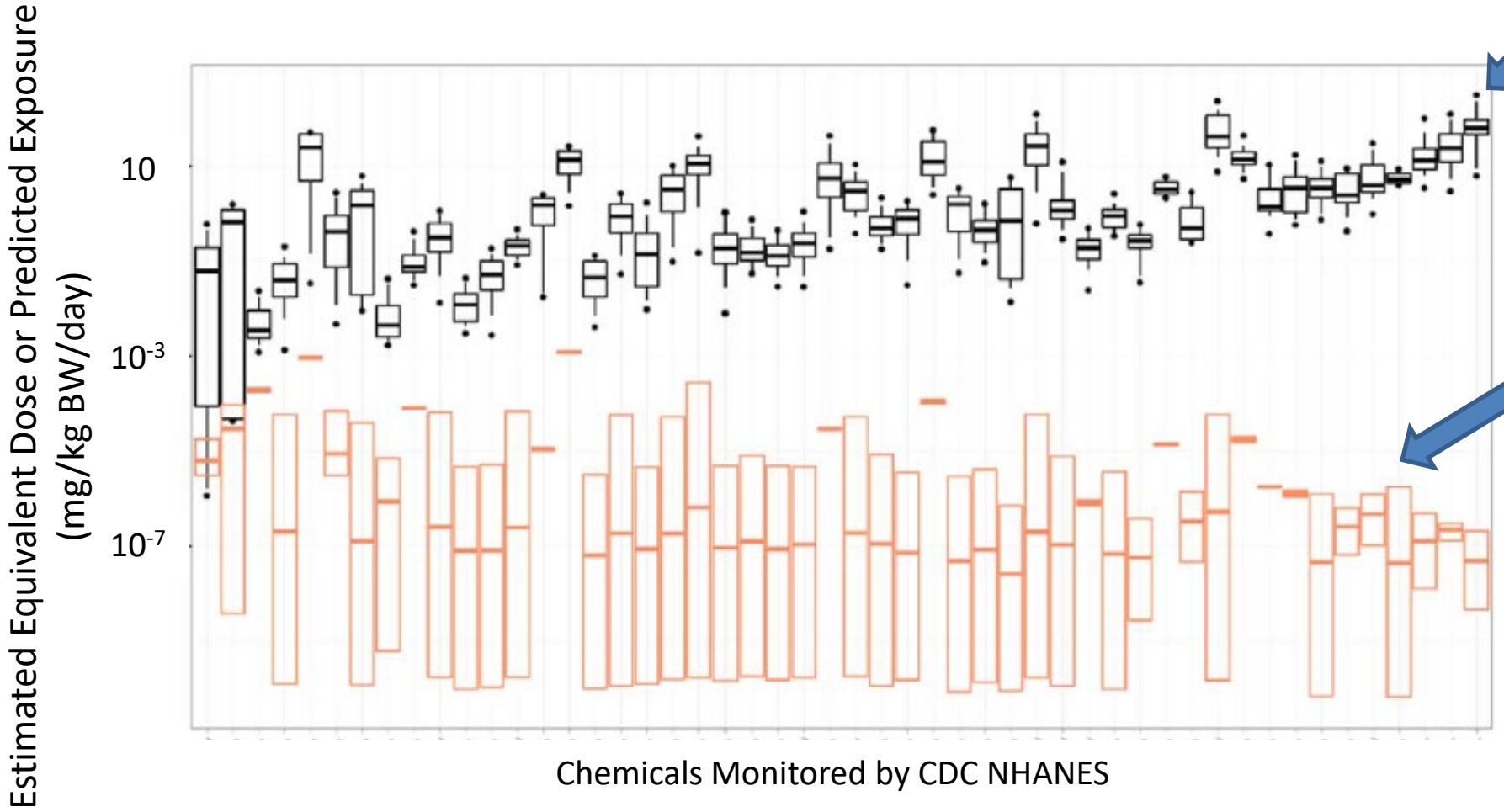
Slide from Risa Sayre

NAMs for Exposure Science

Exposure NAM Class	Description	Traditional Approach	NAM Makes Use of: Measurement	Toxicokinetics	Models	Descriptors	Evaluation	Machine Learning
Measurements	New techniques including screening analyses capable of detecting hundreds of chemicals present in a sample	Targeted (chemical-specific) analyses	-	●	●	●		●
Toxicokinetics	High throughput methods using *in vitro* data to generate chemical-specific models	Analyses based on *in vivo* animal studies	●	-		●		●
HTE Models	Models capable of making predictions for thousands of chemicals	Models requiring detailed, chemical- and scenario-specific information	●	●	-	●		
Chemical Descriptors	Informatic approaches for organizing chemical information in a machine-readable format	Tools targeted at single chemical analyses by humans				-		●
Evaluation	Statistical approaches that use the data from many chemicals to estimate the uncertainty in a prediction for a new chemical	Comparison of model predictions to data on a per chemical basis	●	●	●	●	-	●
Machine Learning	Computer algorithms to identify patterns	Manual Inspection of the Data	●	●		●		-
Prioritization	Integration of exposure and other NAMs to identify chemicals for follow-up study	Expert decision making	●	●	●	●	●	●

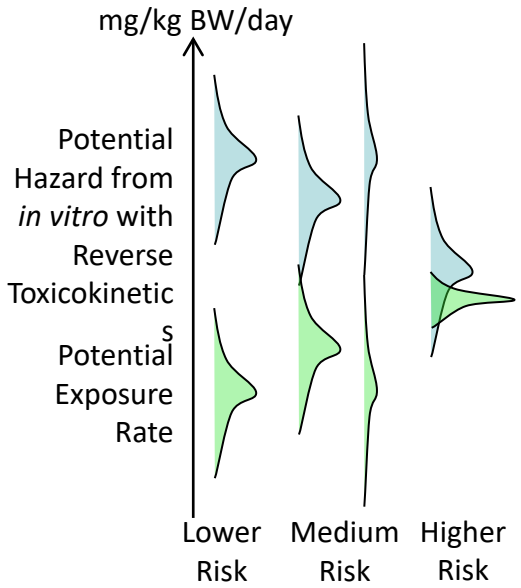
Wambaugh et al., (2019)

Exposure Estimates Allow Chemical Prioritization



High throughput *in vitro* screening can estimate doses needed to cause bioactivity (for example, Wetmore et al., 2015)

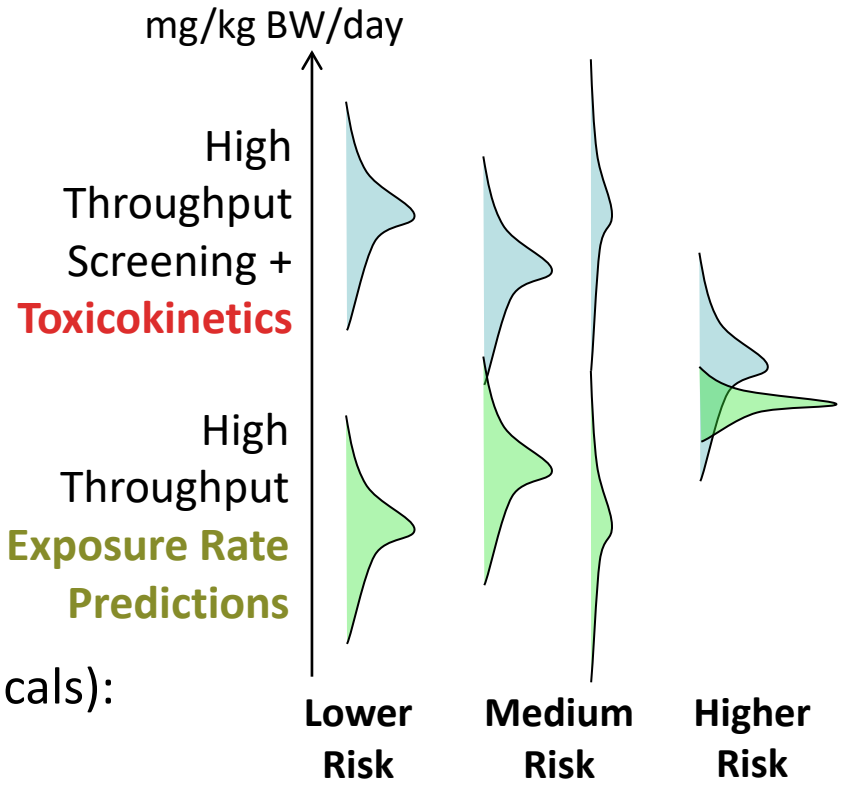
Exposure intake rates can be inferred from biomarkers (for example, Ring et al., 2018)



Ring *et al*. (2017)

Summary

- We need to know chemical hazard, exposure, and toxicokinetics to assess risk posed to the public health
- There are tens of thousands of chemicals in commerce in the environment that lack some of these data
- New approach methodologies (NAMs) are being developed to prioritize these existing and new chemicals for testing
- All data are being made public:
  - The CompTox Chemicals Dashboard (a search engine for chemicals): http://comptox.epa.gov/dashboard
  - R package "httk": https://CRAN.R-project.org/package=httk
  - R package "SEEM3": https://github.com/HumanExposure/SEEM3RPackage
  - R package "SHEDS-HT": https://github.com/HumanExposure/SHEDSHTRPackage

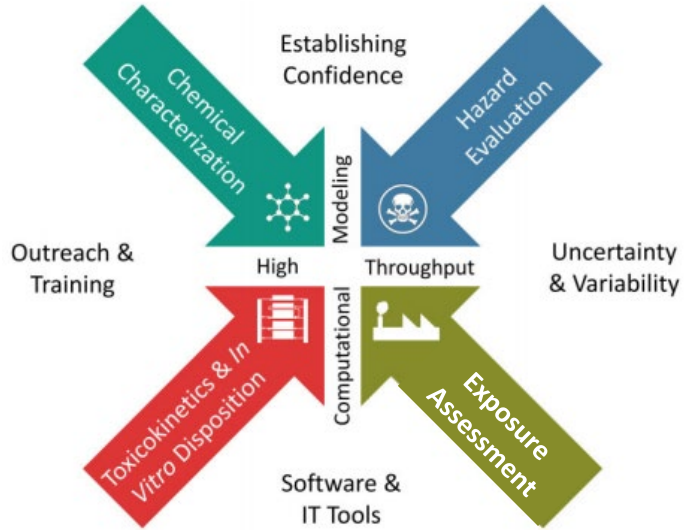


US EPA's ExpoCast Project:
New Approach Methodologies for Exposure Forecasting

*"Investment in **21st century exposure science is now required** to fully realize the potential of the NRC vision for toxicity testing."*
Cohen Hubal (2009)

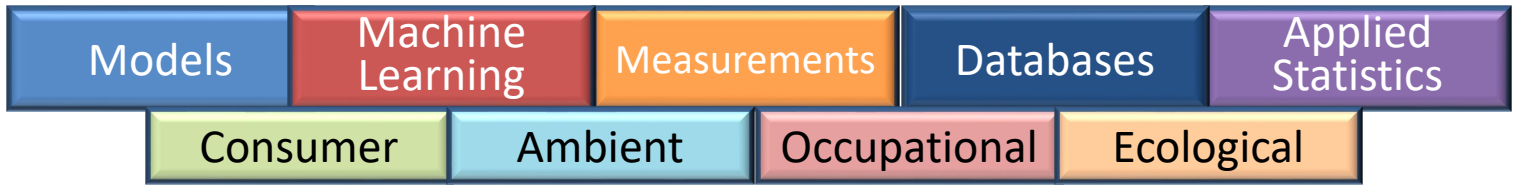
Lovell and Hegstad (2009): "Obama's FY10 Budget Includes Increased Toxicology":

- Funding allows for complementary exposure predictions from ExpoCast, **launched in FY10**
- Predict the impact of chemicals on the human body using data from ToxCast



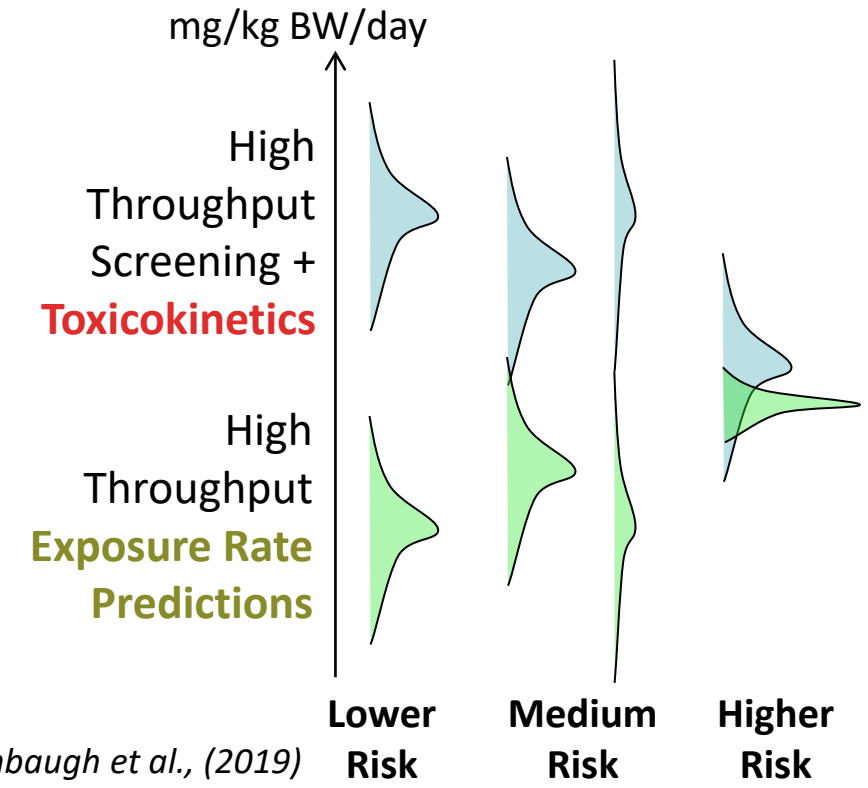
Thomas et al. (2019)

ExpoCast is



Since 2010:

- 45 peer-reviewed publications
- 5 STAR grants awarded
- 3 Federal research contracts (SWRI and Battelle)



Wambaugh et al., (2019)

ExpoCast Project (Exposure Forecasting)

Center for Computational Toxicology and Exposure

Linda Adams
Lucas Albrecht*
Matthew Boyce*
Miyuki Breen*
Alex Chao
Daniel Dawson*
Mike Devito
Alex East*
Lindsay Eddy*
Christopher Eklund
Peter Egeghy
Marina Evans
Alex Fisher*
Rocky Goldsmith
Louis Groff*
Chris Grulke
Colin Guider*
Mike Hughes
Victoria Hull*
Kristin Isaacs
Richard Judson
Jen Korol-Bexell*
Anna Kreutz*
Charles Lowe*
Seth Newton
Katherine Phillips
Paul Price
Tom Purucker
Ann Richard
Caroline Ring
Risa Sayre
Marci Smeltz*
Jon Sobus
Zach Stanfield*
Mike Tornero-Velez
Rusty Thomas
Elin Ulrich
Dan Vallero
Barbara Wetmore
John Wambaugh
Antony Williams

CEMM

Hongwan Li*
Xiaoyu Liu
Zachary Robbins*
Mark Strynar

*Trainees

Collaborators

Arnot Research and Consulting
Jon Arnot
Johnny Westgate
Integrated Laboratory Systems
Xiaoqing Chang
Shannon Bell
National Toxicology Program
Steve Ferguson
Kamel Mansouri
Ramboll
Harvey Clewell
Silent Spring Institute
Robin Dodson
Simulations Plus
Michael Lawless
Southwest Research Institute
Alice Yau
Kristin Favela
Summit Toxicology
Lesa Aylward
Technical University of Denmark
Peter Fantke
Unilever
Beate Nicol
Cecilie Rendal
Ian Sorrell
United States Air Force
Heather Pangburn
Matt Linakis
University of California, Davis
Deborah Bennett
University of Michigan
Olivier Jolliet
University of Texas, Arlington
Hyeong-Moo Shin
University of Nevada
Li Li
University of North Carolina, Chapel Hill
Julia Rager
Marc Serre

United States Environmental Protection Agency

References

Arnot, Jon A., et al. "Screening level risk assessment model for chemical fate and effects in the environment." Environmental science & technology 40.7 (2006): 2316-2323.

Aylward, Lesa L., and Sean M. Hays. "Consideration of dosimetry in evaluation of ToxCast™ data." Journal of Applied Toxicology 31.8 (2011): 741-751.

Breyer, Stephen. Breaking the vicious circle: Toward effective risk regulation. Harvard University Press, 2009

Cohen Hubal, EA, et al. "Advancing internal exposure and physiologically-based toxicokinetic modeling for 21st-century risk assessments." Journal of exposure science & environmental epidemiology (2018).

Collins FS, Gray GM, Bucher JR. Transforming environmental health protection. Science. 2008;319:906–907.

Dionisio, Kathie L., et al. "Exploring consumer exposure pathways and patterns of use for chemicals in the environment." Toxicology reports 2 (2015): 228-237.

Dionisio, Kathie L., et al. "The Chemical and Products Database, a resource for exposure-relevant data on chemicals in consumer products." Scientific data 5 (2018): 180125.

Dix David, et al. "The ToxCast program for prioritizing toxicity testing of environmental chemicals." Toxicol Sci. 2007;95:5–12

Egeghy, P. P., et al. (2012). The exposure data landscape for manufactured chemicals. Science of the Total Environment, 414, 159-166.

Filer, Dayne L., et al. "tcpl: the ToxCast pipeline for high throughput screening data." Bioinformatics 33.4 (2016): 618-620.

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.

Hertzberg, R. P., & Pope, A. J. (2000). high throughput screening: new technology for the 21st century. Current opinion in chemical biology, 4(4), 445-451.

Isaacs, Kristin K., et al. "Consumer product chemical weight fractions from ingredient lists." Journal of Exposure Science and Environmental Epidemiology 28.3 (2018): 216.

Jamei, et al. "The Simcyp® population-based ADME simulator." Expert opinion on drug metabolism & toxicology 2009b;5:211-223

Judson, Richard, et al. "The toxicity data landscape for environmental chemicals." Environmental health perspectives 117.5 (2008): 685-695.

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.

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.

Kavlock, R. J., et al. (2018). Accelerating the pace of chemical risk assessment. Chemical research in toxicology, 31(5), 287-290

MacLeod, Matthew, et al. "The state of multimedia mass-balance modeling in environmental science and decision-making." (2010): 8360-8364.

Mansouri, Kamel, et al. "OPERA models for predicting physicochemical properties and environmental fate endpoints." Journal of cheminformatics 10.1 (2018): 10.

McNally, et al., "PopGen: a virtual human population generator." Toxicology 2014

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.

National Research Council. Exposure Science in the 21st Century: a Vision and a Strategy. National Academies Press, 2012.

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 19.4 (2017): 1063-1074.

Phillips, Katherine A., et al. "Suspect screening analysis of chemicals in consumer products." Environmental science & technology 52.5 (2018): 3125-3135.

Ring, Caroline L., et al. "Identifying populations sensitive to environmental chemicals by simulating toxicokinetic variability." Environment International 106 (2017): 105-118.

Ring, Caroline L., et al. "Consensus modeling of median chemical intake for the US population based on predictions of exposure pathways." Environmental science & technology 53.2 (2018): 719-732.

Rotroff, Daniel M., et al. "Incorporating human dosimetry and exposure into high throughput in vitro toxicity screening." Toxicological Sciences 117.2 (2010): 348-358

Schmidt, Charles W. "TOX 21: new dimensions of toxicity testing." Environmental health perspectives 117.8 (2009): A348.

Shibata, Yoshihiro, et al. "Prediction of hepatic clearance and availability by cryopreserved human hepatocytes: an application of serum incubation method." Drug Metabolism and disposition 30.8 (2002): 892-896.

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.

Sipes, Nisha S., et al. "An intuitive approach for predicting potential human health risk with the Tox21 10k library." Environmental science & technology 51.18 (2017): 10786-10796.

US Congress. "Frank R. Lautenberg Chemical Safety for the 21st Century Act." (2016).

U.S. E.P.A. (2018) "A Working Approach for Identifying Potential Candidate Chemicals for Prioritization." https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/identifying-existing-chemicals-prioritization-under-tsca

U.S. G.A.O.. "Toxic substances: EPA has increased efforts to assess and control chemicals but could strengthen its approach." (2013).

Wallace, Lance A., 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 43.2 (1987): 290-307.

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 147.1 (2015): 55-67.

Wambaugh, John F., et al. "Evaluating in vitro-in vivo extrapolation of toxicokinetics." Toxicological Sciences 163.1 (2018): 152-169.

Wambaugh, John F., et al. "Assessing Toxicokinetic Uncertainty and Variability in Risk Prioritization" Toxicological Sciences (2019), *in press*

Wambaugh, John F., et al. "New Approach Methodologies for Exposure Science." Current Opinion in Toxicology (2019).

Wang, Ying-Hong. "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 (2010): 1094-1104.

Waters, Nigel J., et al. "Validation of a rapid equilibrium dialysis approach for the measurement of plasma protein binding." Journal of pharmaceutical sciences 97.10 (2008): 4586-4595.

Wetmore, Barbara A., et al. "Integration of dosimetry, exposure and high throughput screening data in chemical toxicity assessment." *Tox. Sciences* (2012)

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.