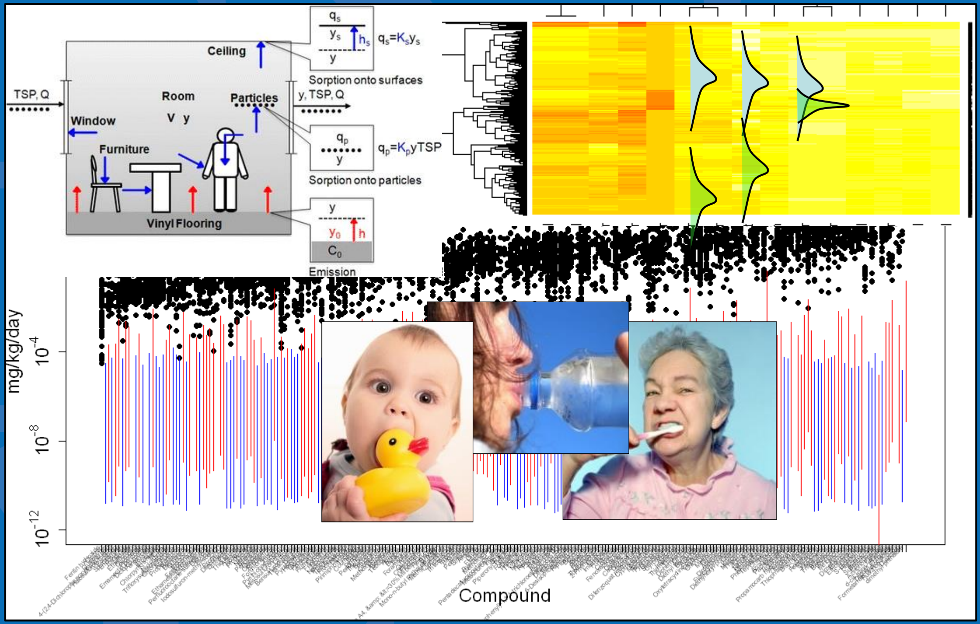
Updates on EPA's High-Throughput Exposure Forecast (ExpoCast) Research project,

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



CompTox Community of Practice, November 20, 2014

Introduction

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

- ExpoCast is an EPA ORD initiative to develop the necessary approaches and tools for rapidly predicting exposure for thousands of chemicals (Cohen-Hubal, *et al.*, 2010)
- **Proof of Concept (First Generation Analysis):** Used off-the-shelf high throughput exposure models – simple description of near field exposure predicted more than existing HT models (Wambaugh *et al.,* 2013)

Environmental Fate and Transport

Consumer Use and Indoor Exposure

“All cases are unique, and very similar to others.”— T.S. Eliot

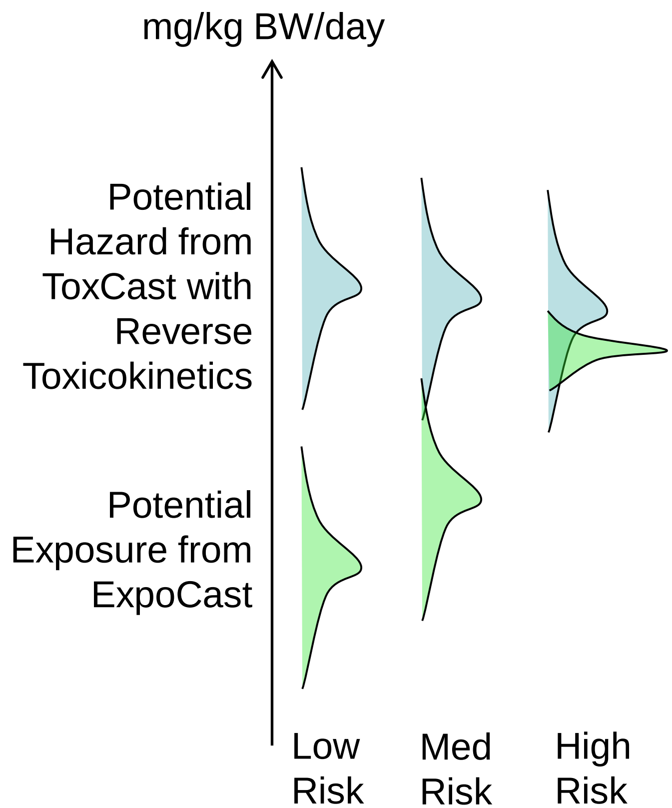
EPA United States Environmental Protection Agency

Risk-based Prioritization Requires Exposure

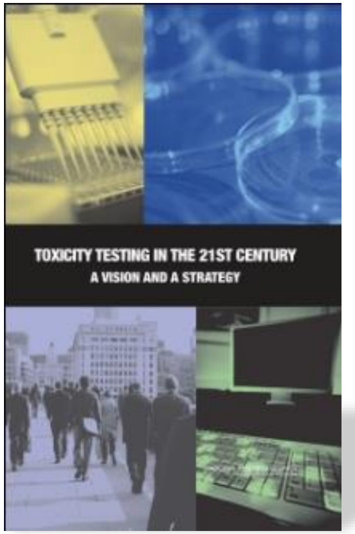
- **Tox21/ToxCast**: Examining thousands of chemicals using high throughput screening assays to identify *in vitro* concentrations that perturb biological pathways (Schmidt, 2009)

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

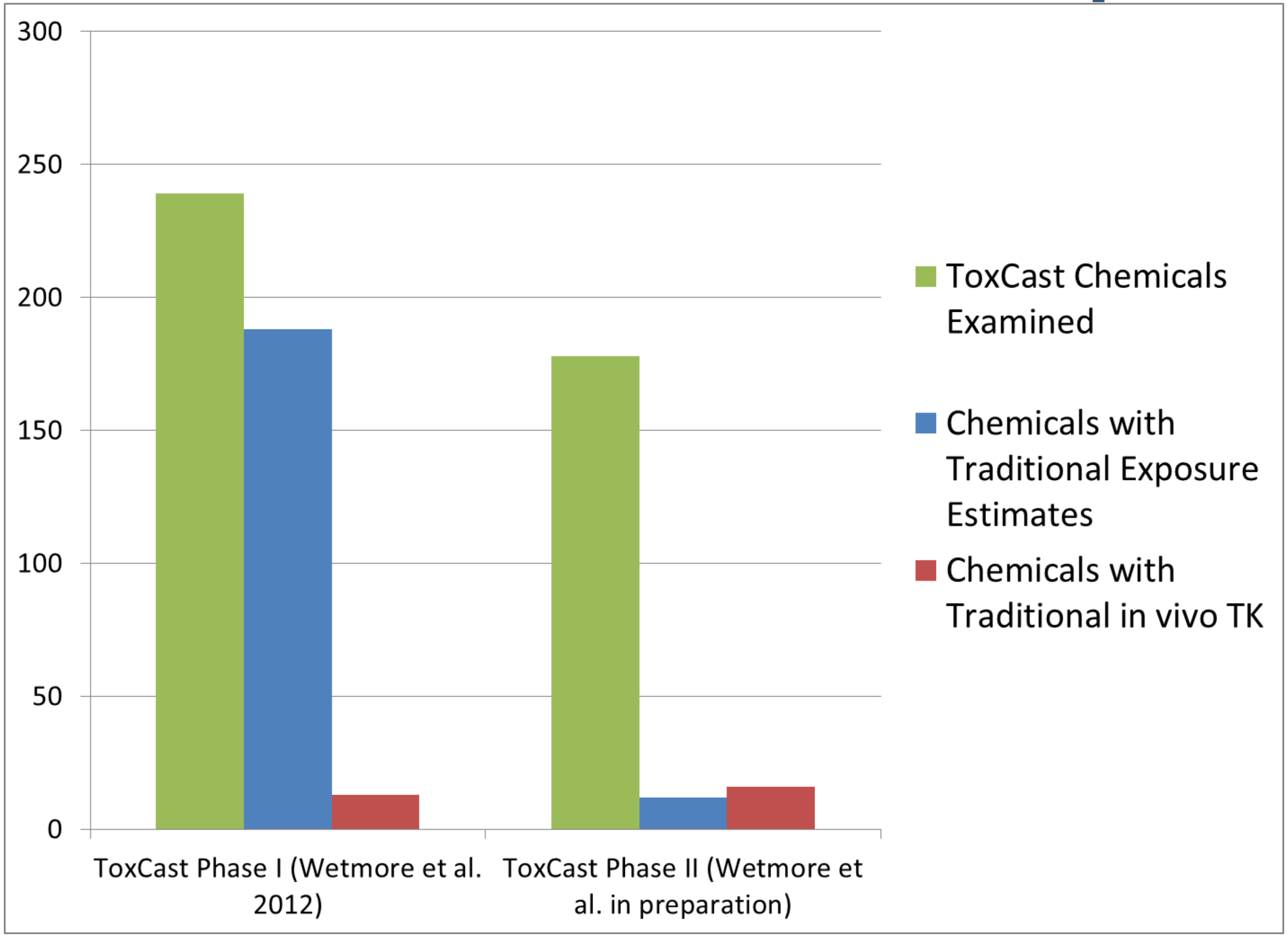
- These doses can then be directly compared with exposure rates, ***where available***



e.g. Judson *et al.*, (2011)

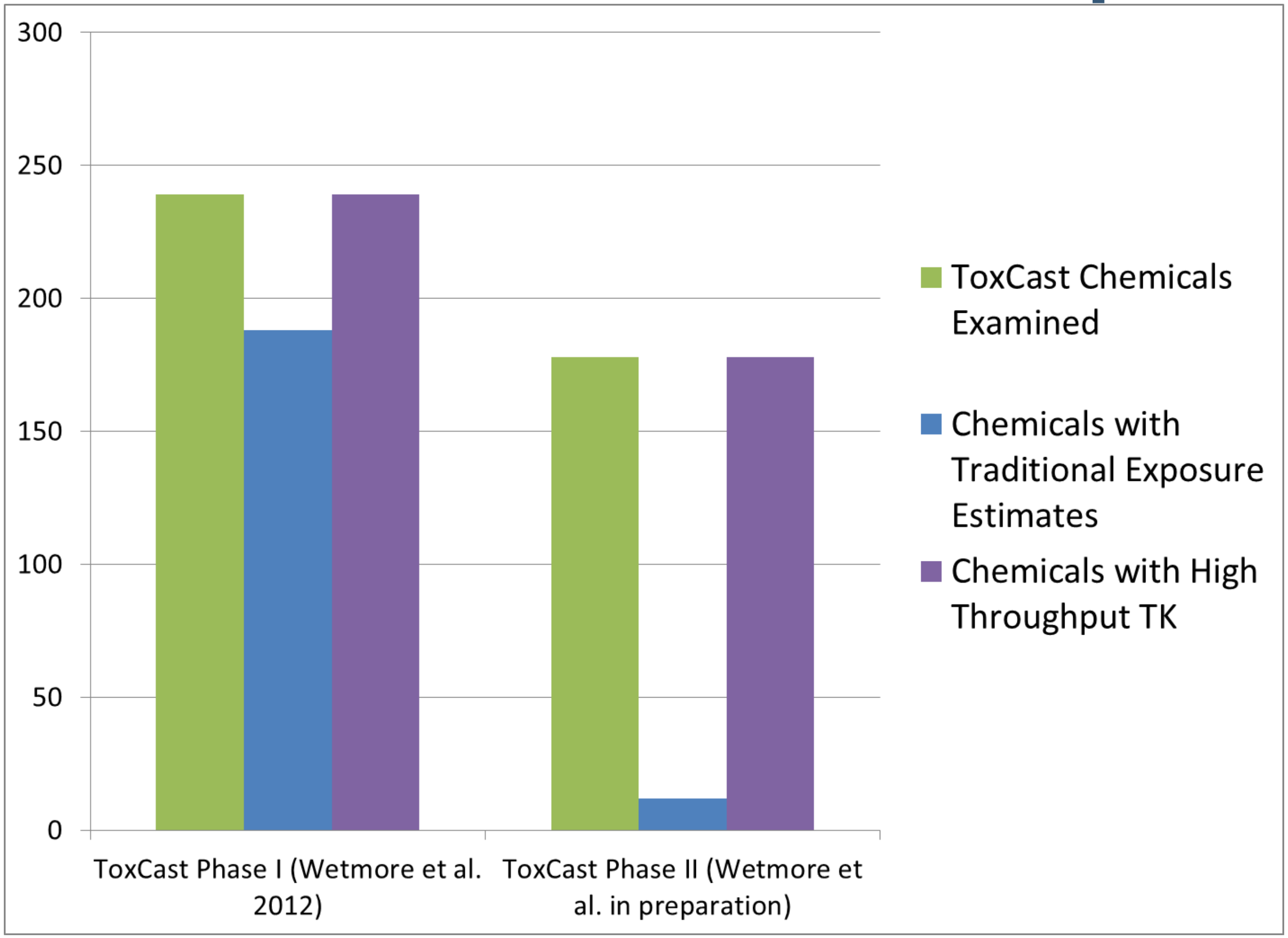


In Vitro Bioactivity, *In Vivo* Toxicokinetics, and Exposure



- Studies like Wetmore et al. (2012),addressed the need for toxicokinetic data

In Vitro Bioactivity, *In Vitro* Toxicokinetics, and Exposure



- As in Egeghy *et al.* (2012), there is a paucity of data for providing context to HTS data

Exposure Science in the 21st Century

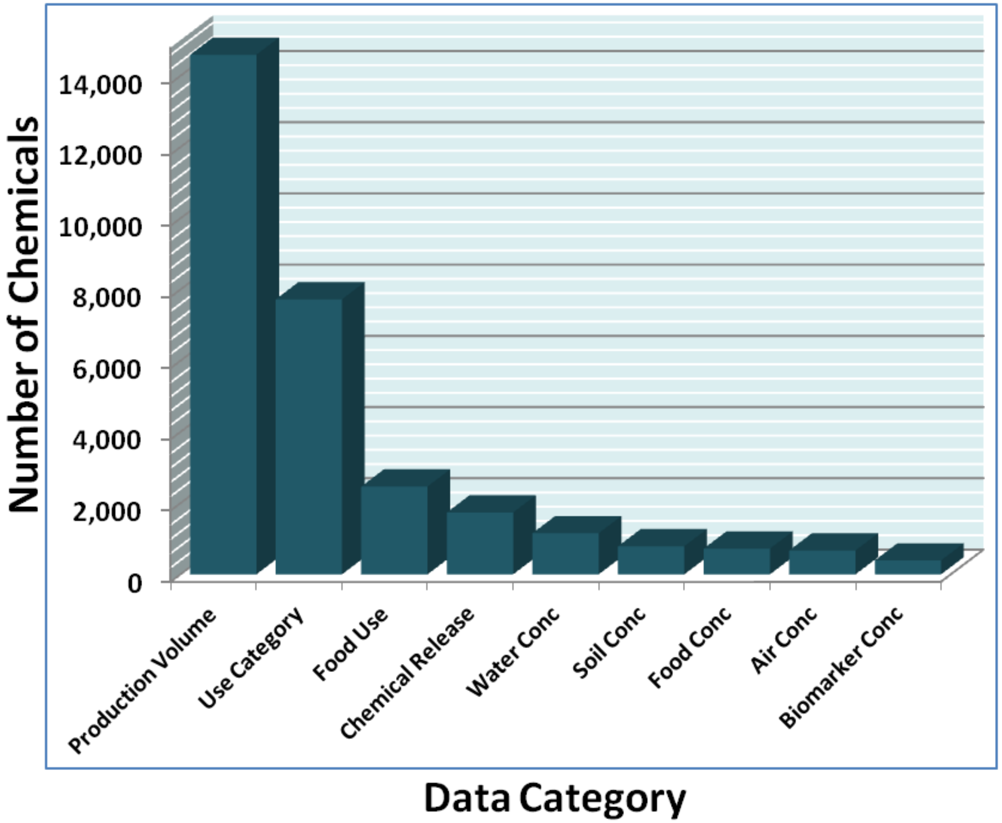
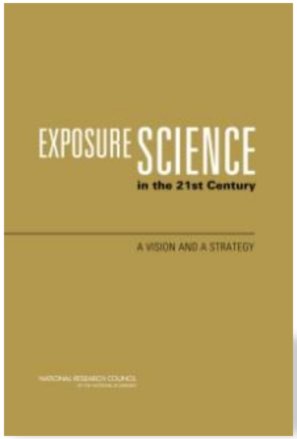


Figure from Egeghy et al. (2012),
"The exposure data landscape for manufactured chemicals"

- 2012 NRC report:



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

Exposure Space

United States Environmental Protection Agency

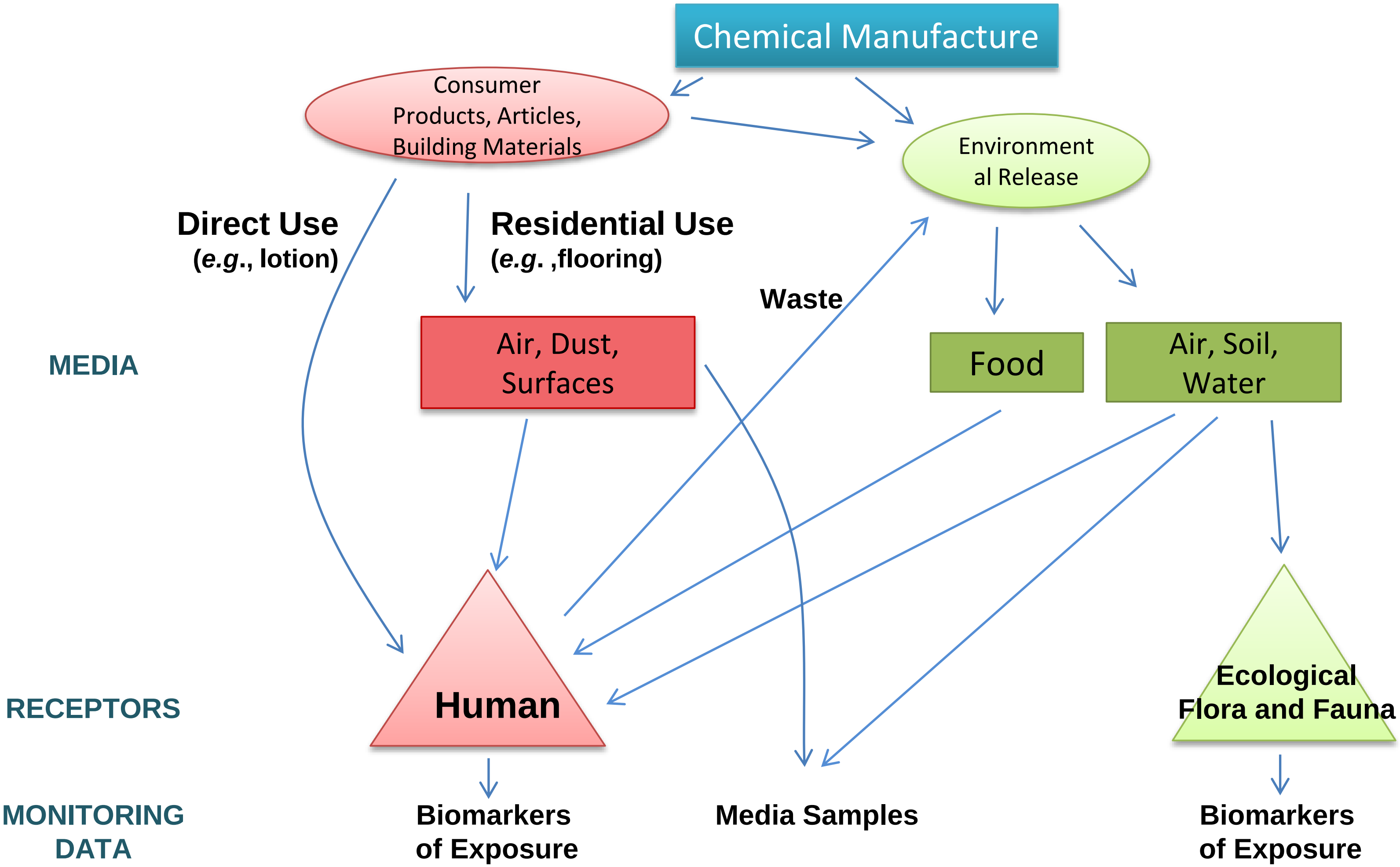


Figure from Kristin Isaacs

Exposure Pathways

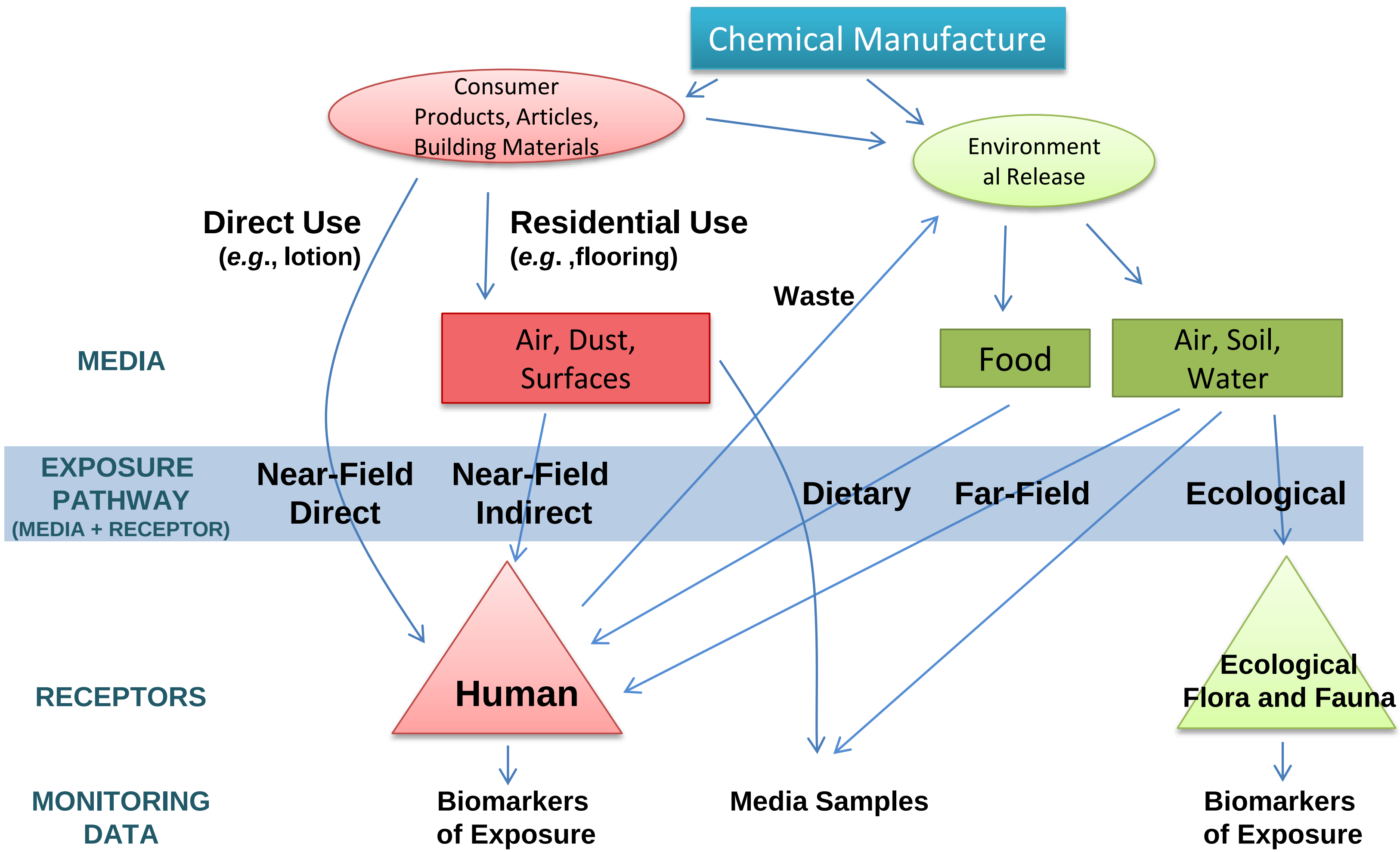


Figure from Kristin Isaacs

Forward Modeling of Exposure Pathways

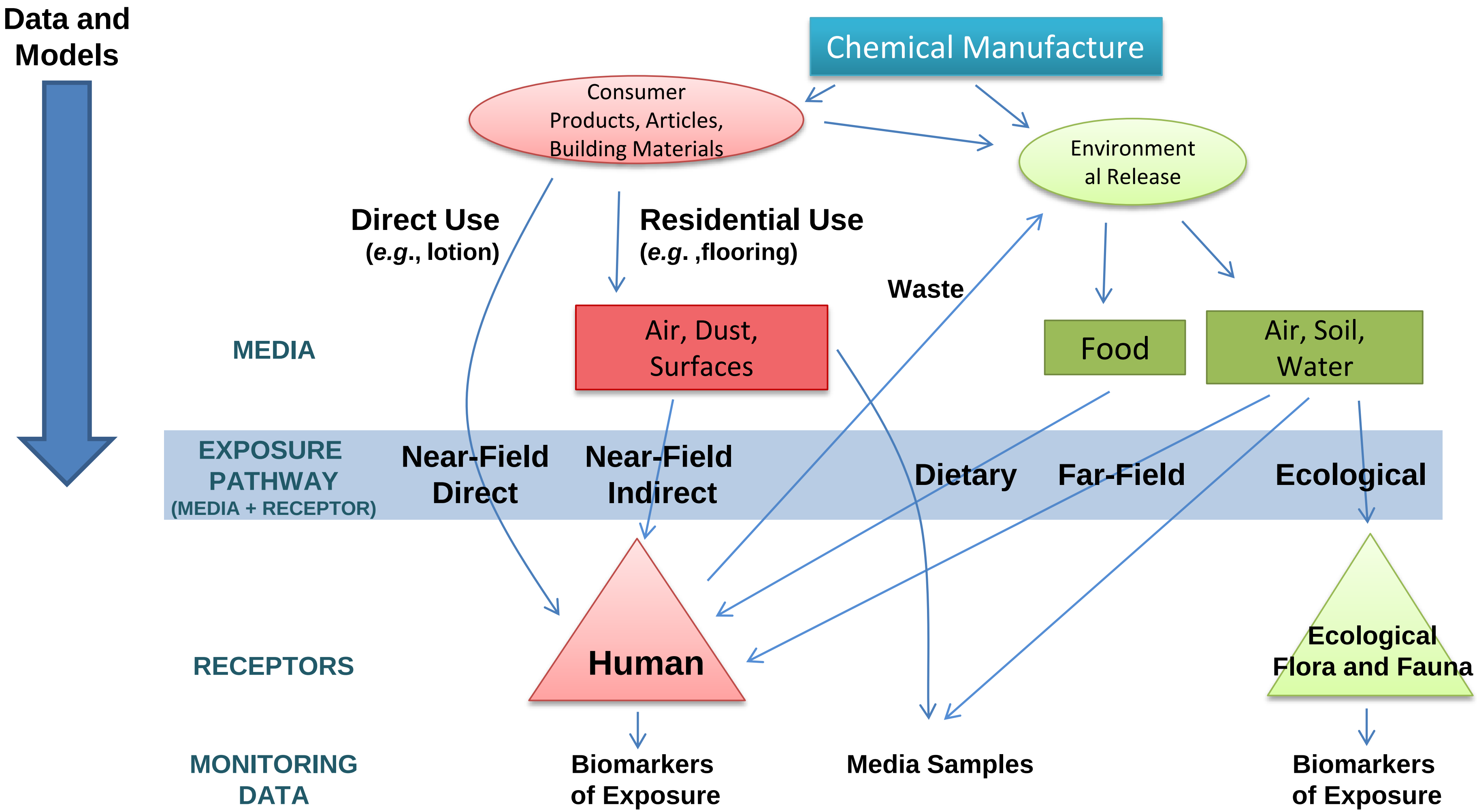
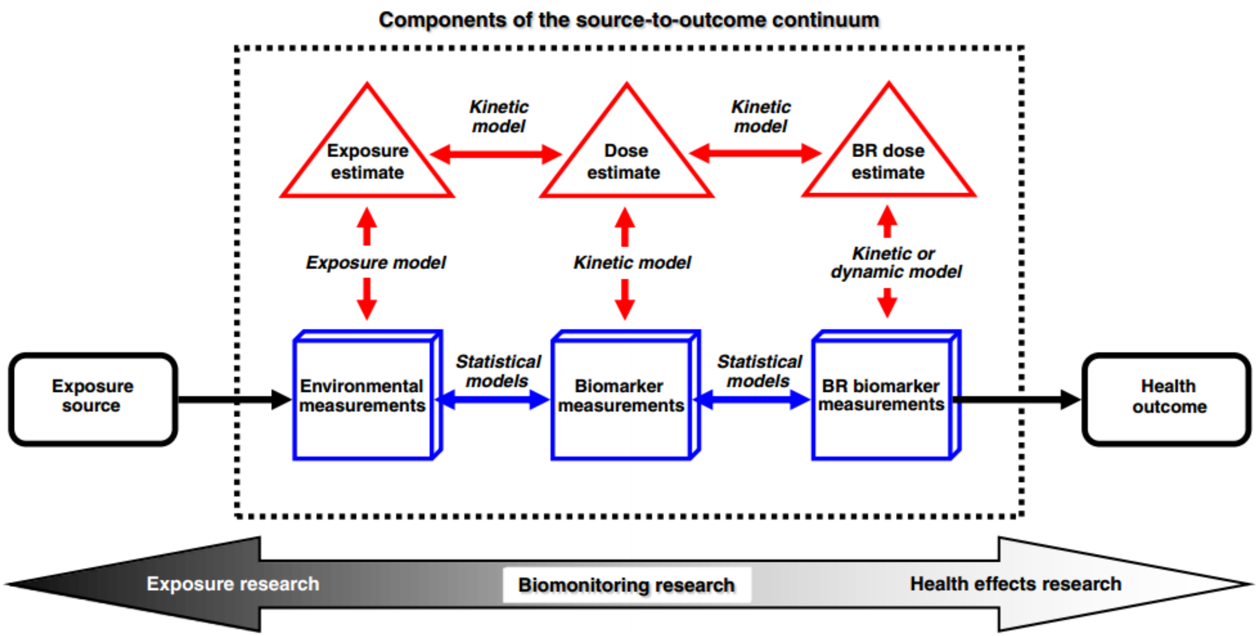


Figure from Kristin Isaacs

Forward Predicting Exposure



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

Sobus *et al.* (2011)

Inference of Exposure Pathways

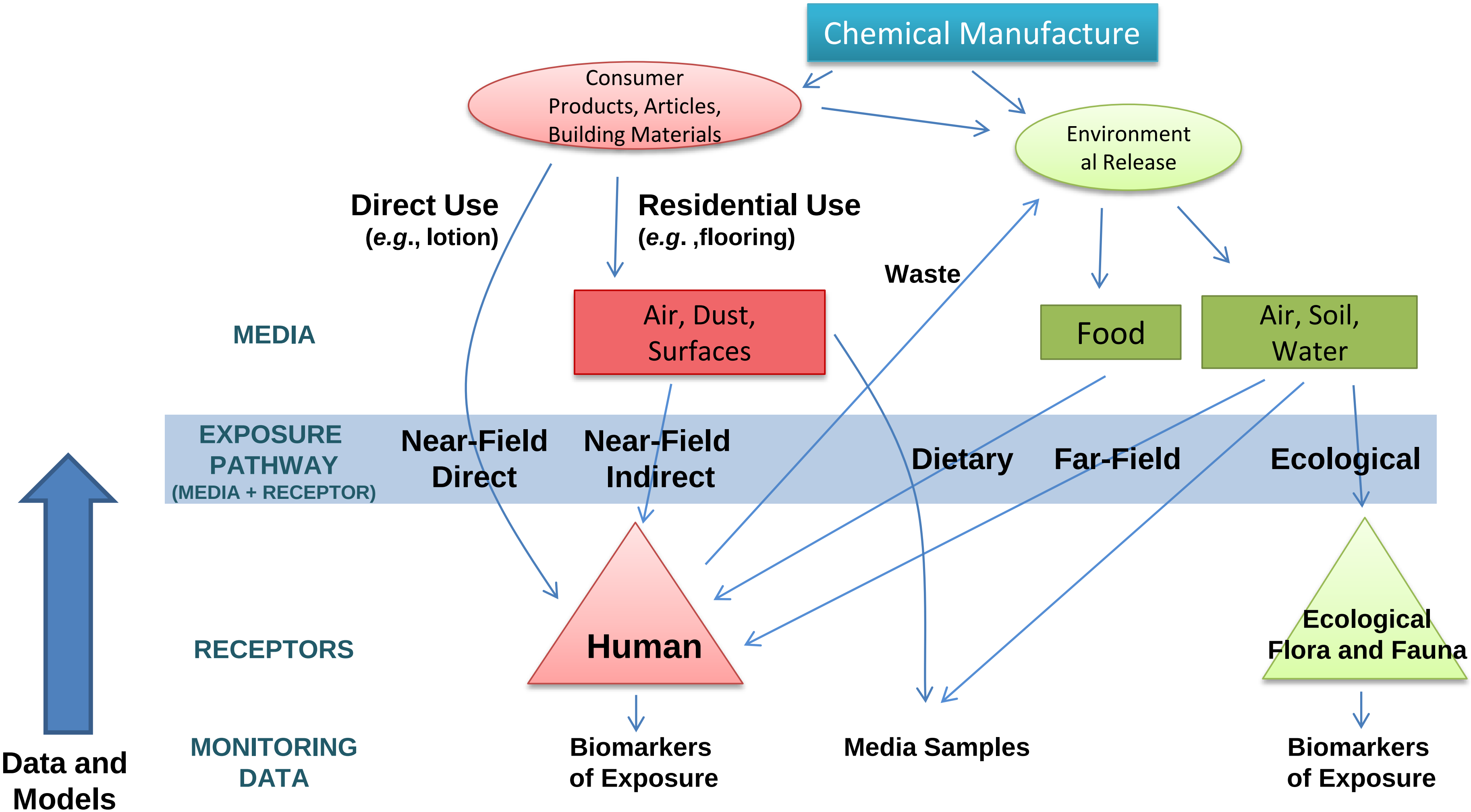
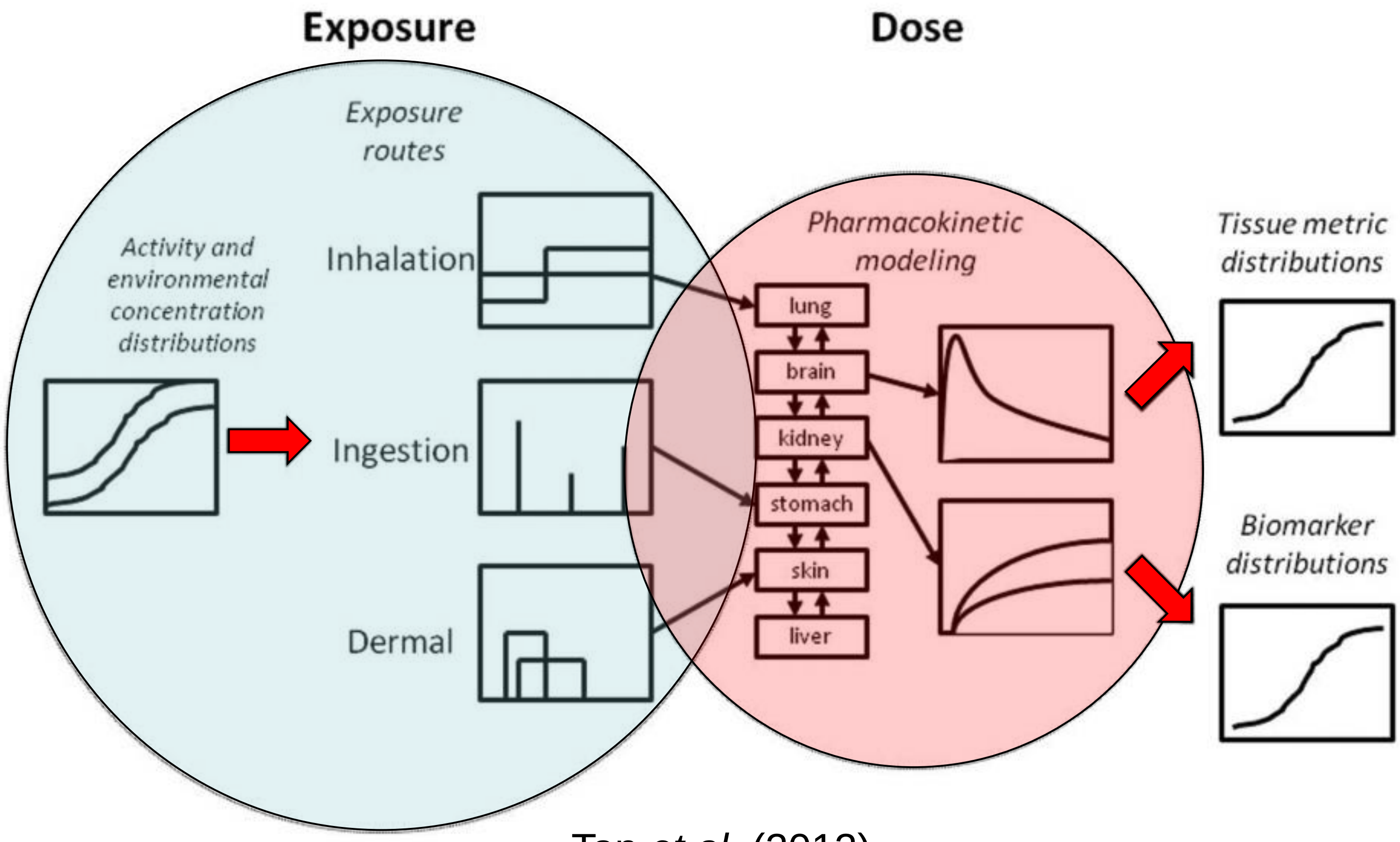


Figure from Kristin Isaacs

Inferring Exposure



Tan *et al.* (2012)

Evaluation of Forward Predictions with Inferred Exposure

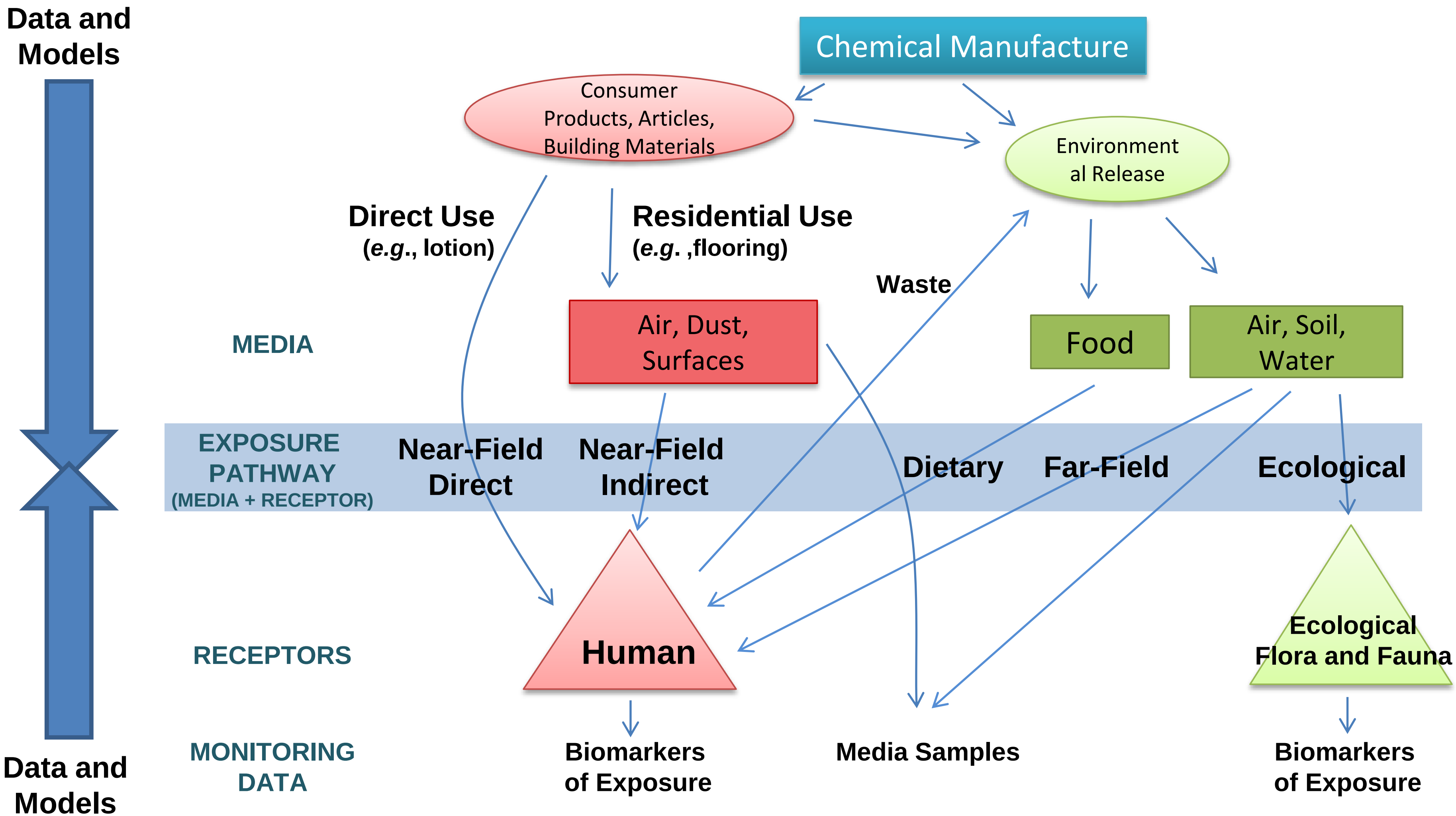
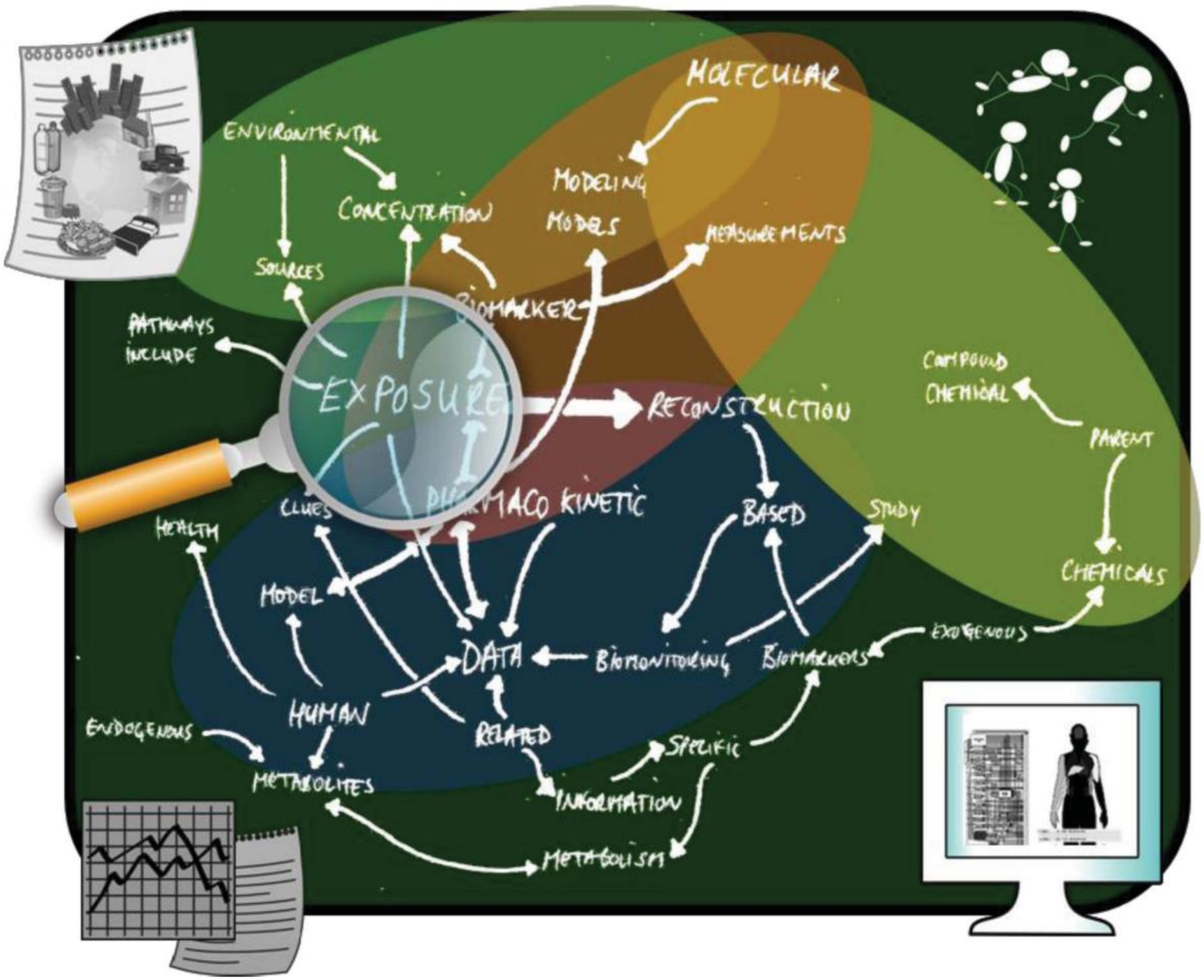


Figure from Kristin Isaacs

Investigating Exposure to Environmental Chemicals



Tan *et al.* (2012):

A cartoon illustrating the relation of different factors and knowledge domains in the exposure reconstruction process. This cartoon is generated using key terms in this review and their semantic/lexical relationships using the visual analysis of IBM's www.many-eyes.com Phrasenet analysis.

EPA
United States
Environmental Protection
Agency

Exposure Detective Work

- Sobus et al. (2011): Use a mix of empirical and mechanistic models
- Empirical models can be as simple as “rule of thumb”, *i.e.* heuristics of exposure

86° Sunny—still one day closer to death, though

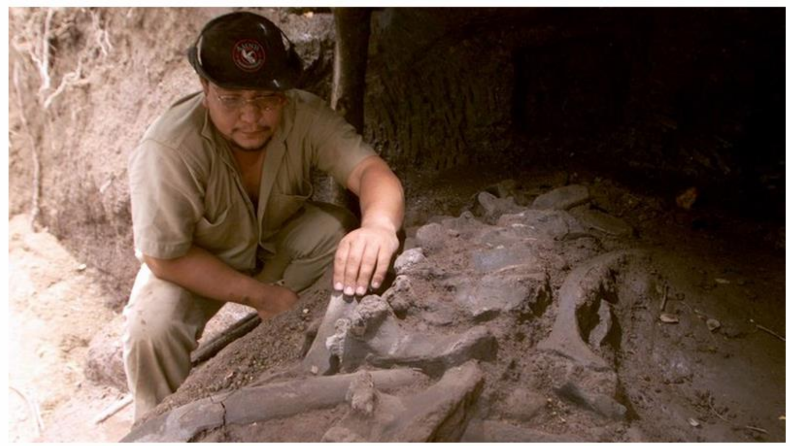
the ONION®
America's Finest News Source

VIDEO • POLITICS • SPORTS • SCIENCE/TECH • LOCAL • ENTERTAINMENT •

Paleontologists Determine Dinosaurs Were Killed By Someone They Trusted

NEWS IN BRIEF • Science & Technology • Death • Science • Animals • ISSUE 50•46 • Nov 17, 2014

Share on Facebook 45.3K Share on Twitter 1.9K 426



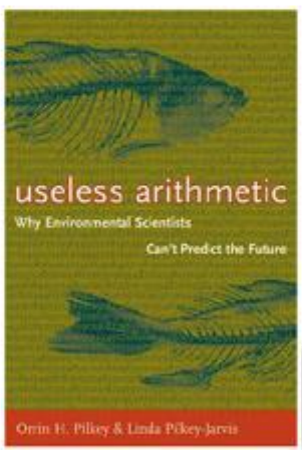
BERKELEY, CA—Citing compelling fossil evidence that the prehistoric species died suddenly and treacherously, paleontologists at the University of California, Berkeley announced Monday that dinosaurs were almost certainly killed by someone they trusted. “Our findings indicate that someone, we don’t know who, spent at least 150 million years gaining the confidence of dinosaurs before abruptly betraying them and taking their lives near the end of the Cretaceous Era,” said lead researcher Professor Janet Bower, adding that dinosaurs likely had an innately innocent and unsuspecting nature that this individual could exploit to get within easy striking distance. “The distribution and condition of dinosaur bones strongly suggests that these creatures died without a struggle and that they had been caught totally off-guard by an individual they naively considered a friend. Those that had time to regard their killer were no doubt

How to Make Good Forecasts

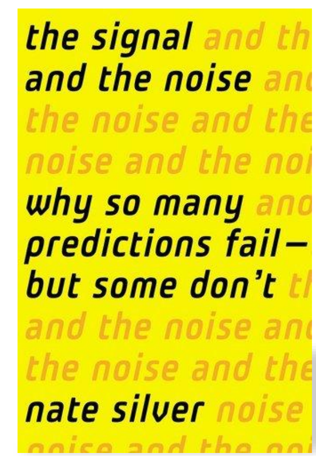
1) Think probabilistically (especially, Bayesian): We use an approach that evaluates model performance systematically across as many chemicals (and chemistries) as possible

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

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



Orrin Pilkey &
Olinda Pilkey-Jarvis (2007)



Nate Silver (2012)

Exposure Forecasting (ExpoCast)

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

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

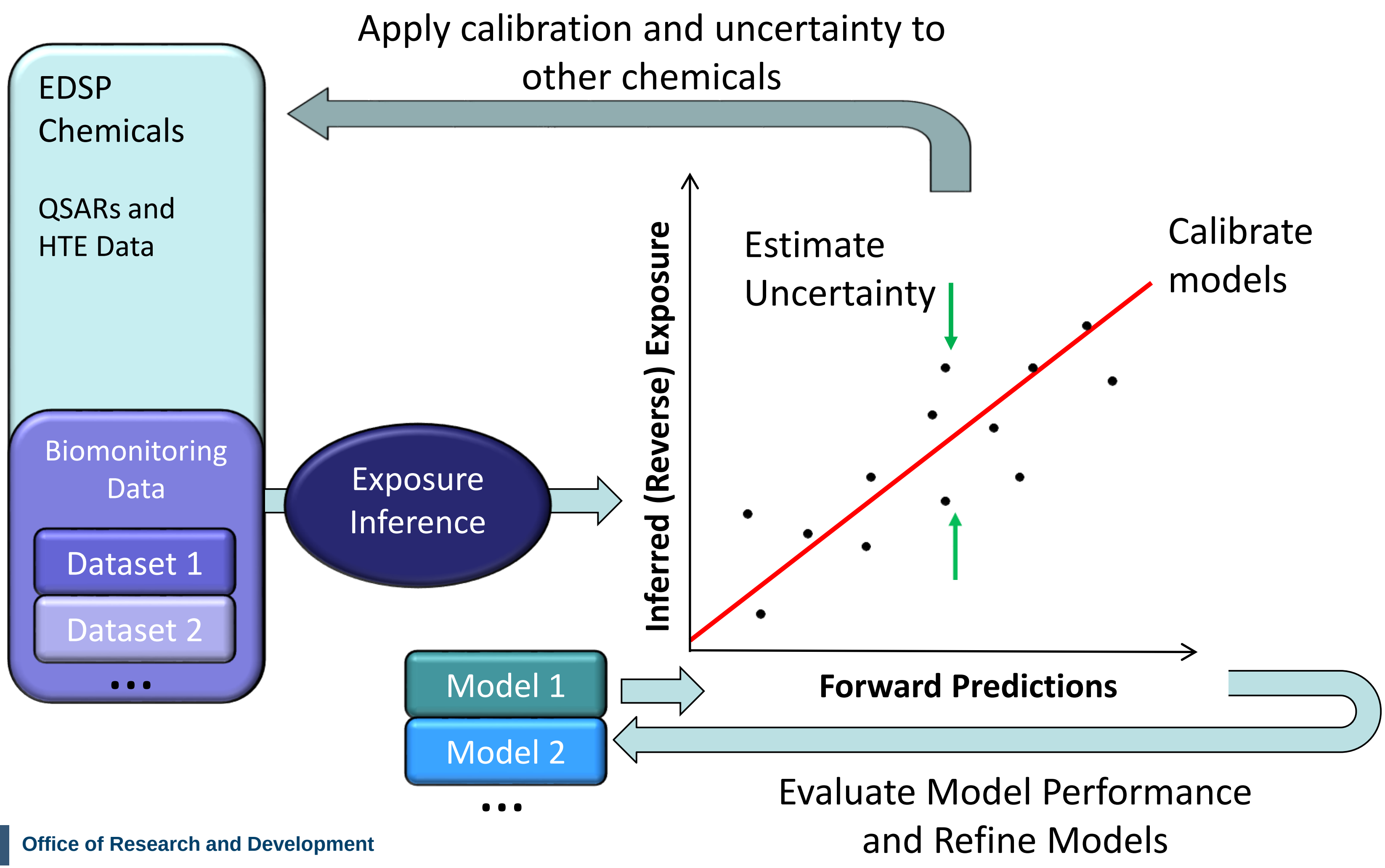
In Nate Silver's terminology:
a ***prediction*** is a specific statement
a ***forecast*** is a probabilistic statement

Wikipedia (statistics): "when information is transferred across time, often to specific points in time, the process is known as forecasting"

Systematic Empirical Evaluation of Models (SEEM)

- There are four basic steps in the SEEM framework
  1. Forward prediction of exposures, which involves model curation and parameterization
  2. Inference of exposures from monitoring data
  3. Systematic evaluation and calibration of the predictions against the inferred exposures
  4. Extrapolation of the calibrated model predictions and estimated uncertainty to chemicals with no monitoring data.
- To achieve these aims the SEEM framework used Bayesian formalism and multivariate, linear regression for demonstrating and evaluating predictive ability

Illustration of the SEEM Framework



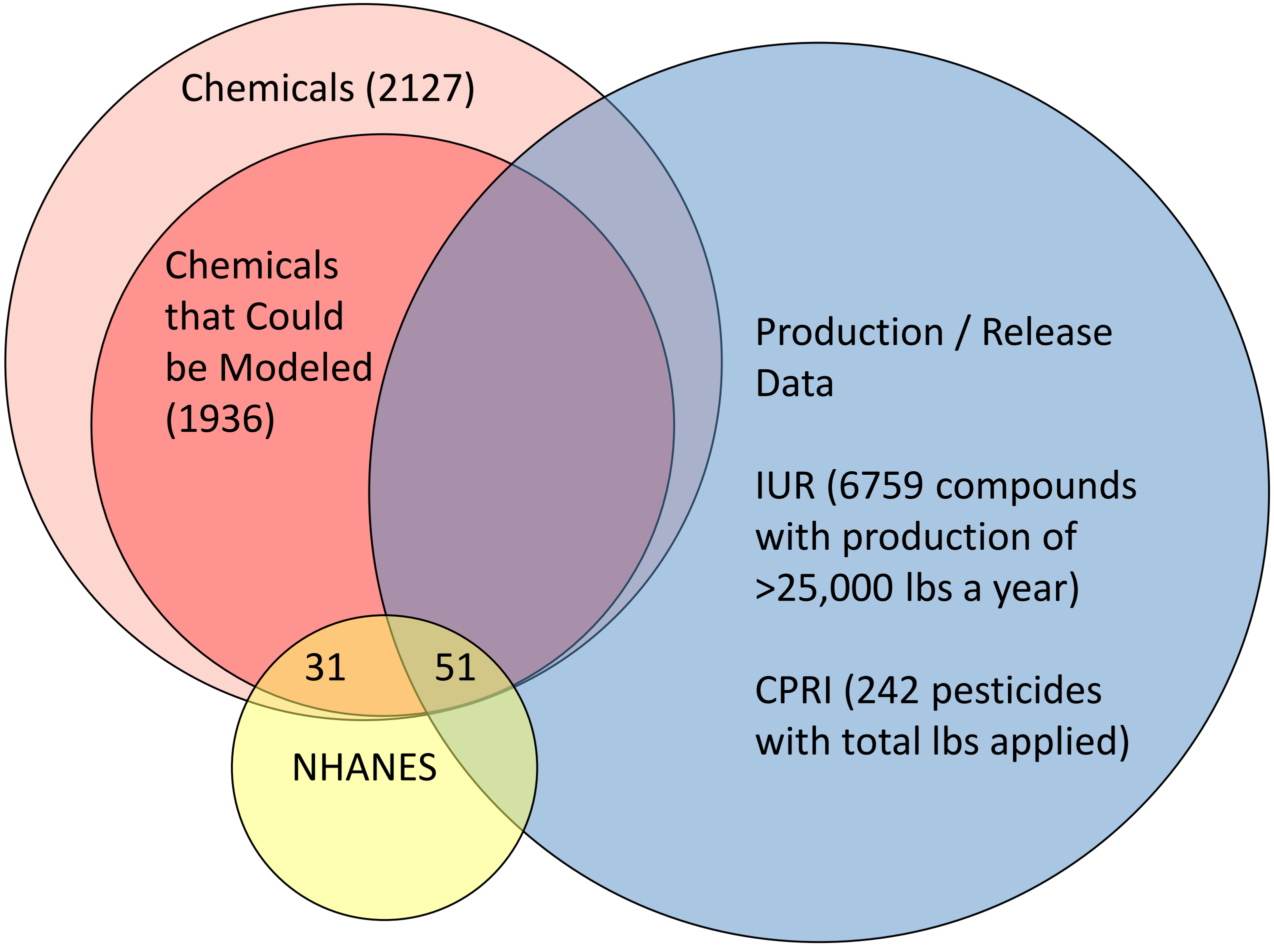
Goals for High Throughput Exposure

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

Image from Thinkstock

Data Availability for Evaluating Model Predictions

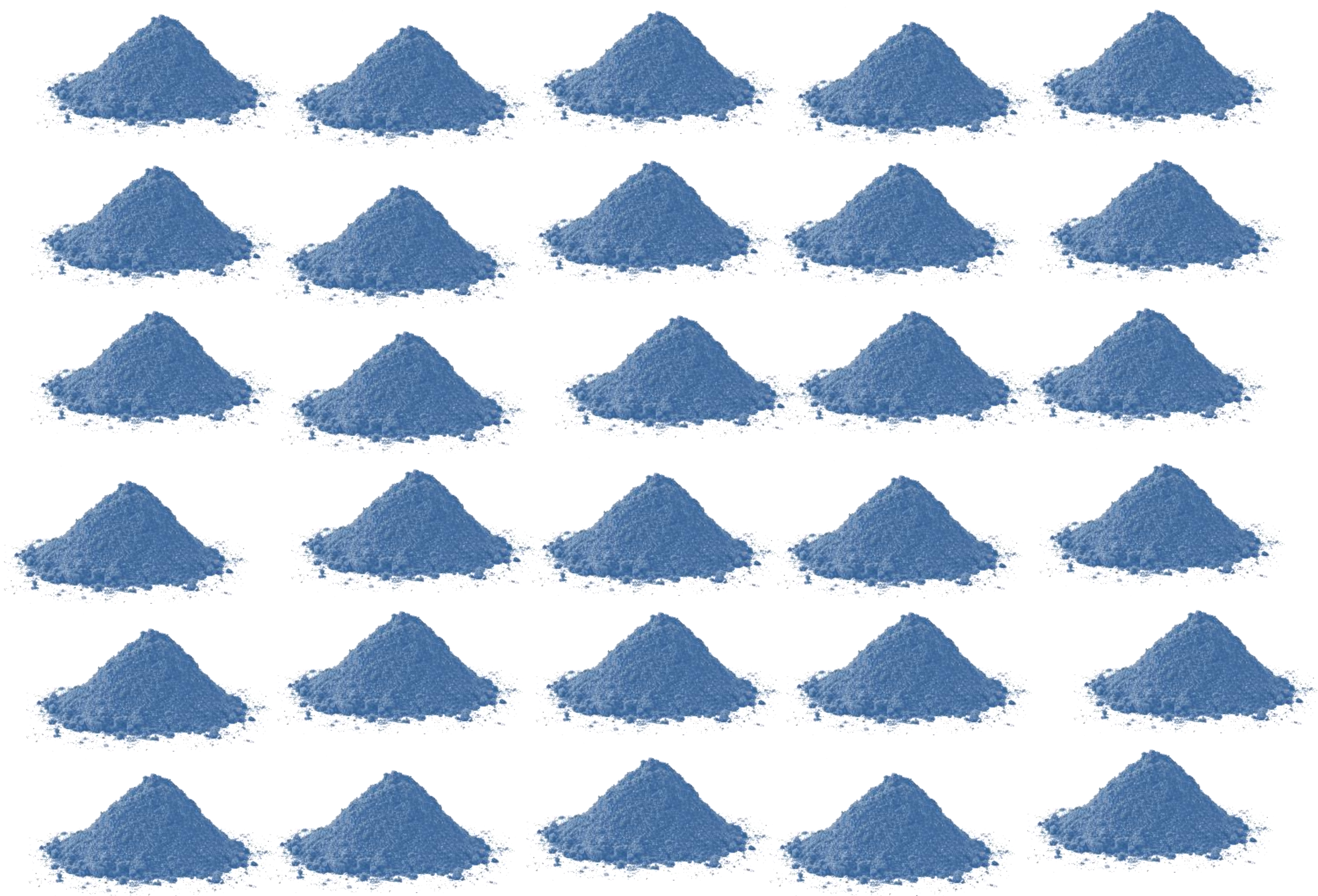
- Currently we use the CDC NHANES urine data
- Many chemicals had median conc. below the limit of detection (LoD)
  - Most chemicals >LoD not high production volume
- 106 chemicals inferred from urine to date
- Dozens more expected with serum/blood model



Wambaugh *et al.* (2013)

EPA United States Environmental Protection Agency

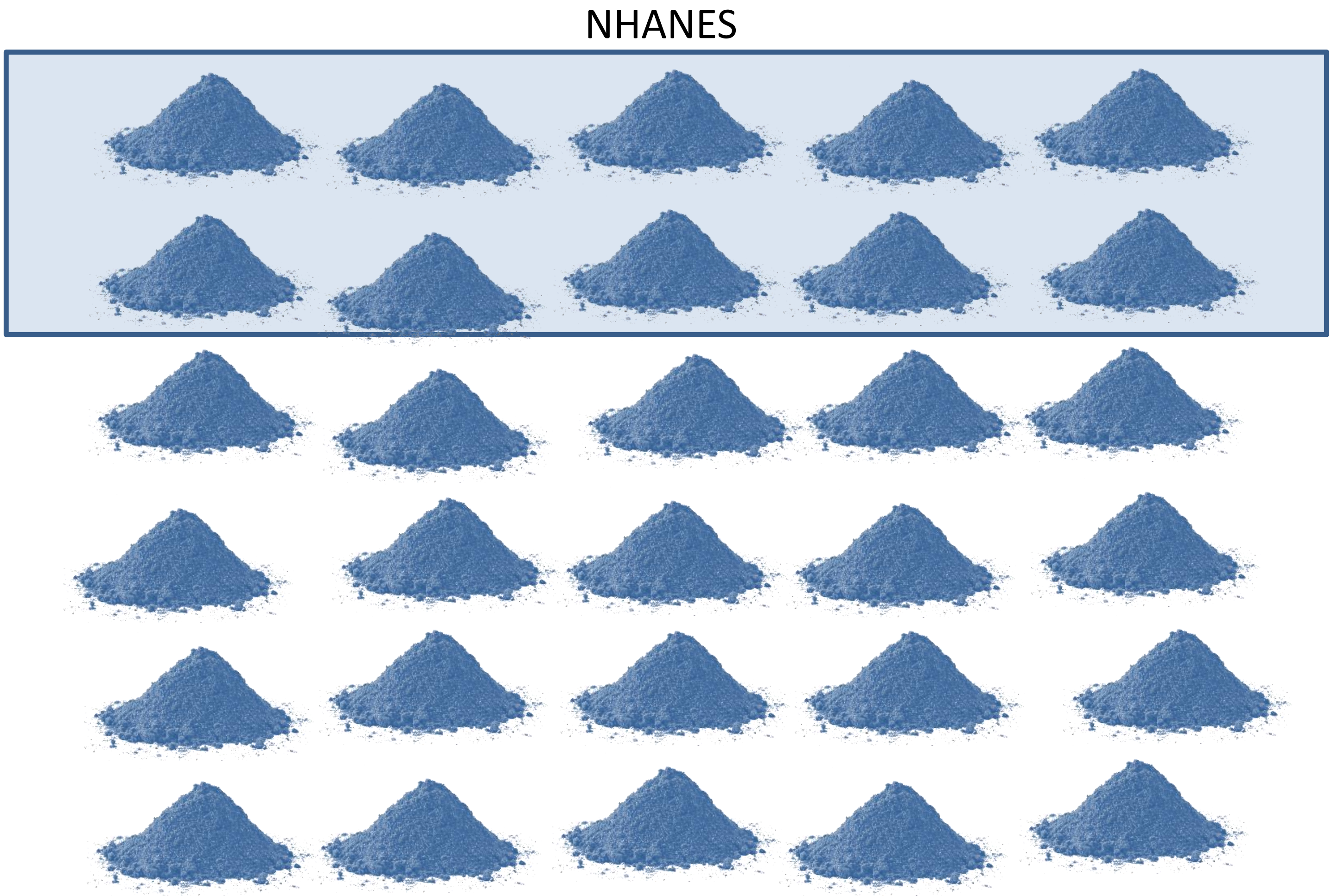
Applying ExpoCast



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

Images from Thinkstock

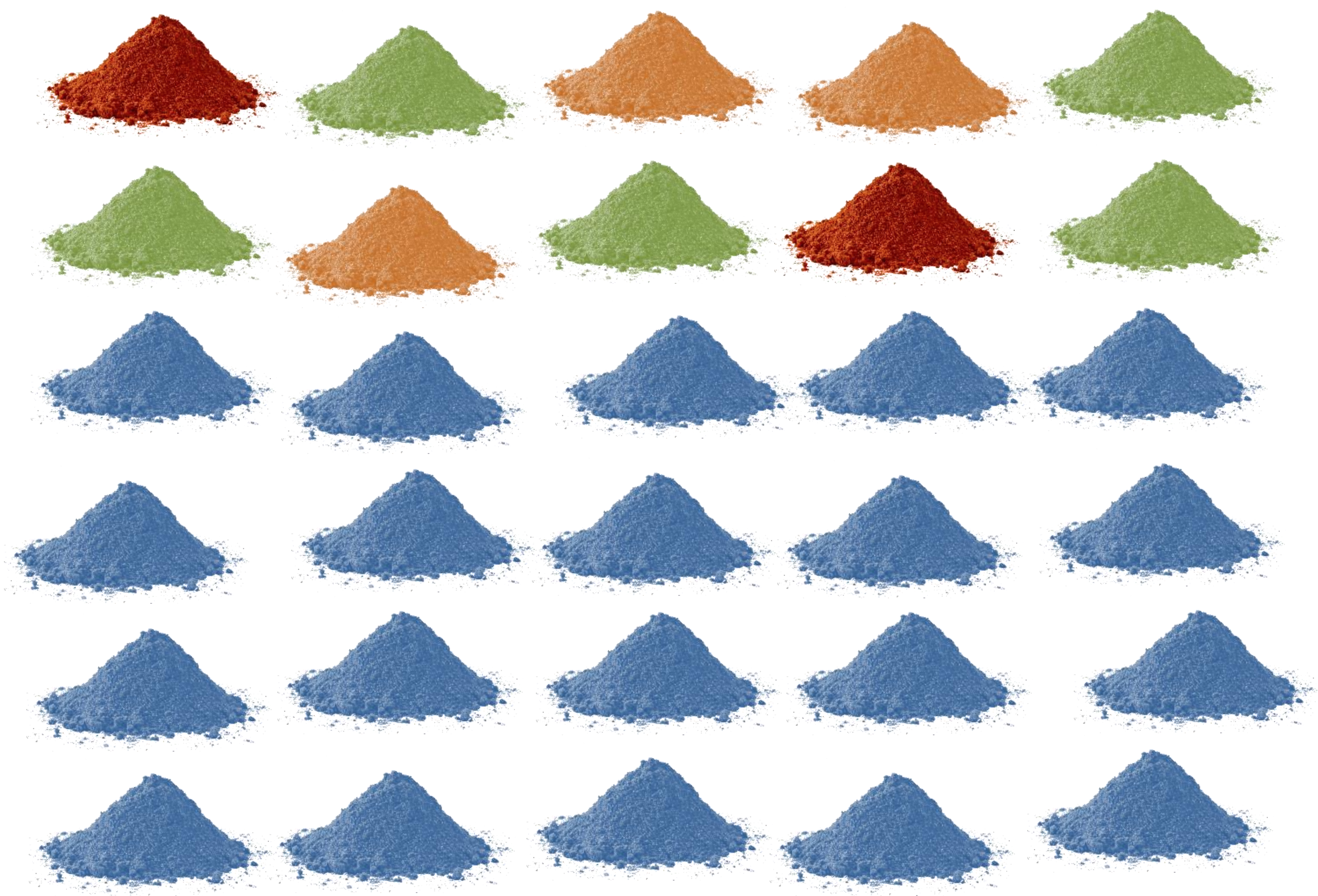
Applying ExpoCast



- The CDC targets some chemicals for exposure biomonitoring

Images from Thinkstock

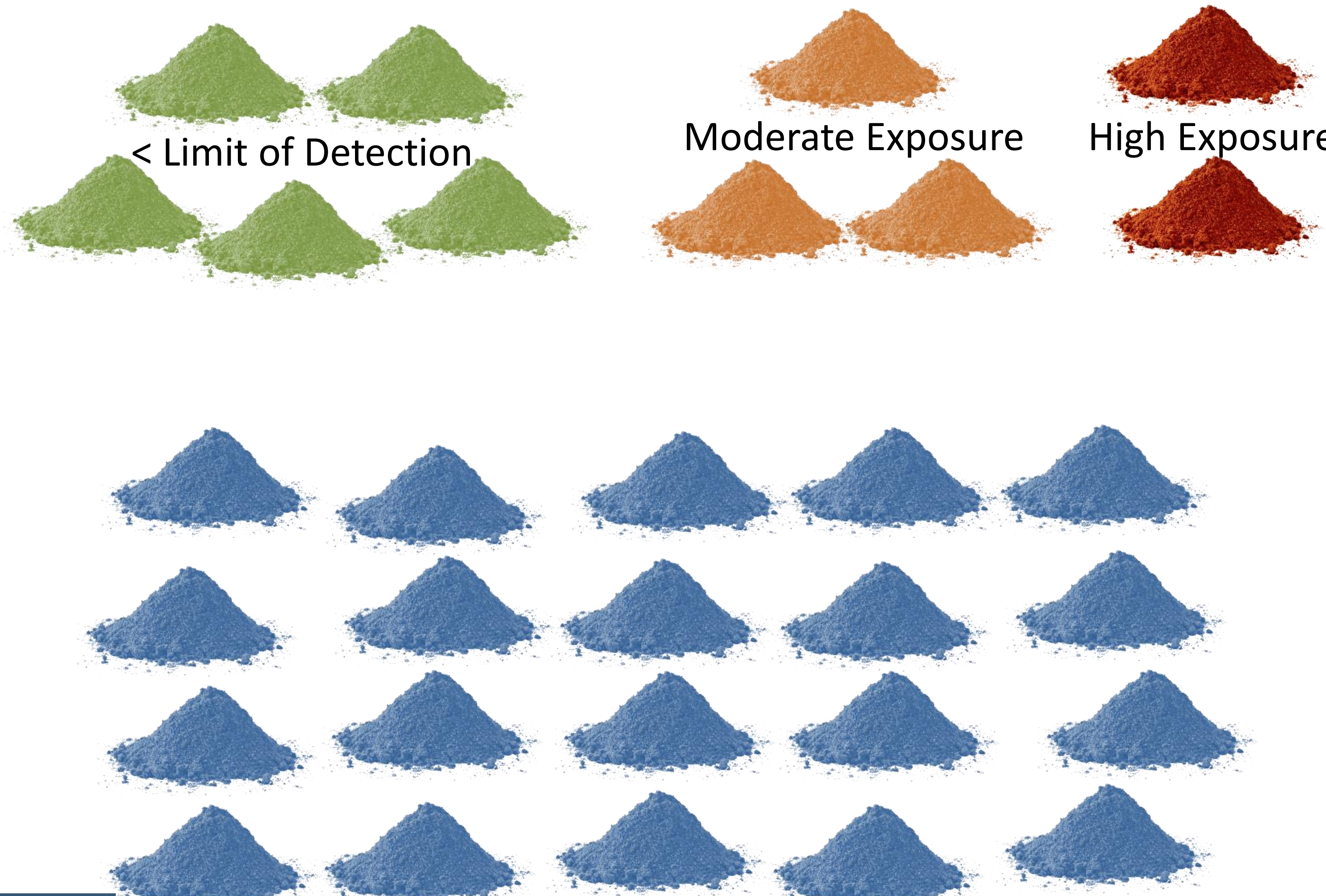
Applying ExpoCast



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

Images from Thinkstock

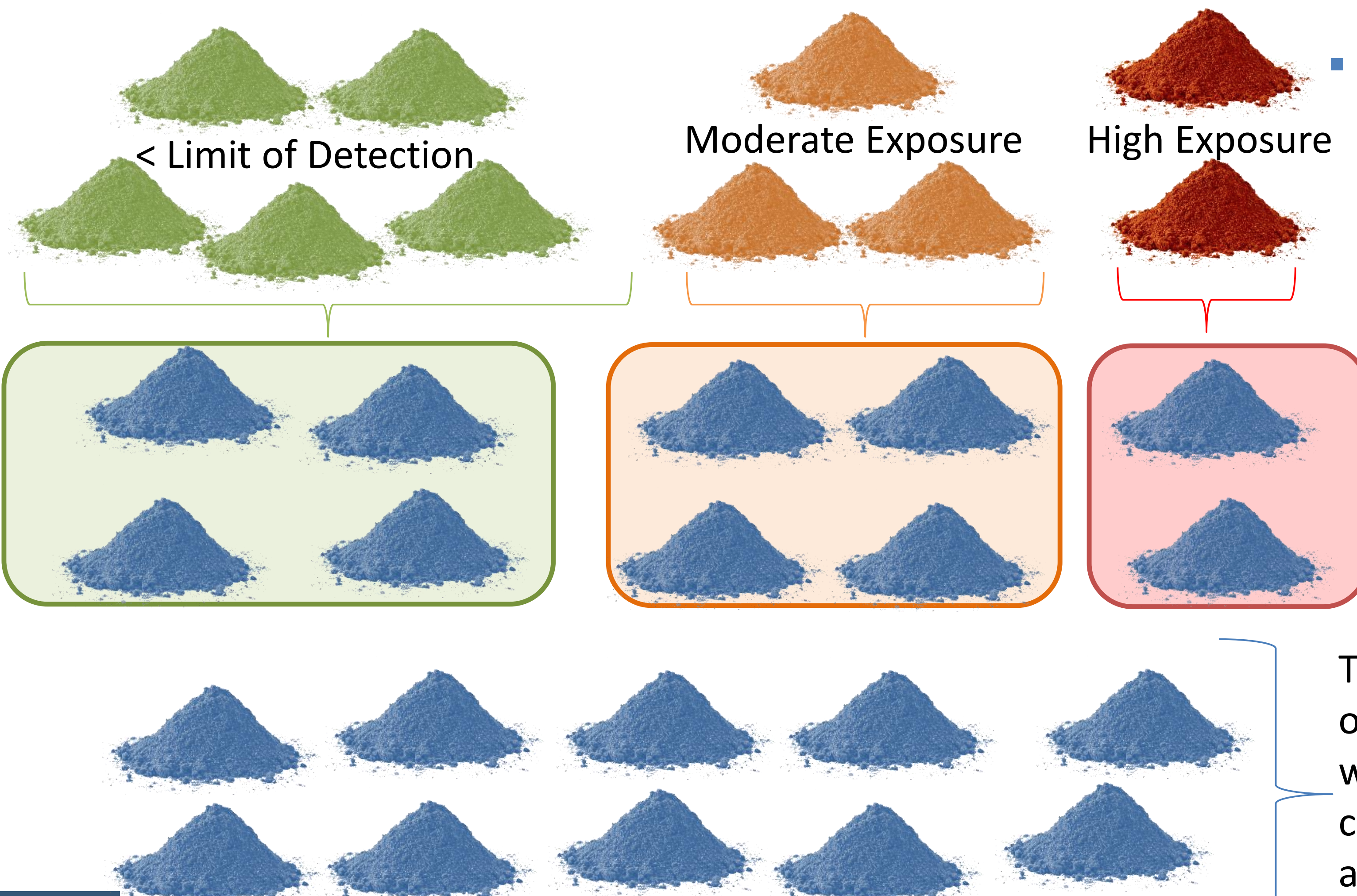
Applying ExpoCast



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

Images from Thinkstock

Applying ExpoCast



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

There will still be other chemicals without characteristics that are predictive of NHANES chemicals

Images from Thinkstock

NHANES is Much More than a Chemical Survey

- Separate evaluations can be done for various demographics

Urinary Bisphenol A (2,2-*bis*[4-Hydroxyphenyl] propane)

Geometric mean and selected percentiles of urine concentrations (in µg/L) for the U.S. populatio
and Nutrition Examination Survey.

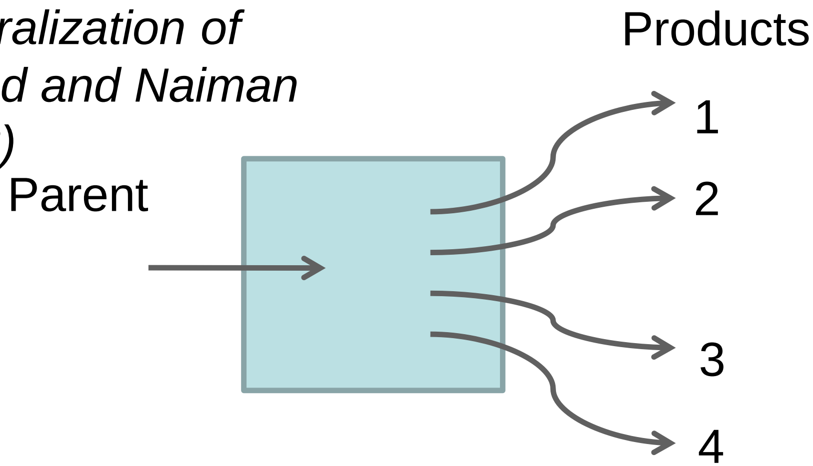
	Survey years	Geometric mean (95% conf. interval)	Selected percentiles (95% confidence interval)		
			50th	75th	90th
Total	03-04	2.64 (2.38-2.94)	2.80 (2.50-3.10)	5.50 (5.00-6.20)	10.6 (9.40-
	05-06	1.90 (1.79-2.02)	2.00 (1.90-2.00)	3.70 (3.50-3.90)	7.00 (6.40-
	07-08	2.08 (1.92-2.26)	2.10 (1.90-2.30)	4.10 (3.60-4.60)	7.70 (6.80-
Age group					
6-11 years	03-04	3.55 (2.95-4.29)	3.80 (2.70-5.00)	6.90 (6.00-8.30)	12.6 (9.50-
	05-06	2.86 (2.52-3.24)	2.70 (2.30-2.90)	5.00 (4.40-5.80)	13.5 (9.30-
	07-08	2.46 (2.20-2.75)	2.40 (1.90-3.00)	4.50 (3.70-5.50)	7.00 (6.30-
12-19 years	03-04	3.74 (3.31-4.22)	4.30 (3.60-4.60)	7.80 (6.50-9.00)	13.5 (11.8-
	05-06	2.42 (2.18-2.68)	2.40 (2.10-2.70)	4.30 (3.90-5.20)	8.40 (6.50-
	07-08	2.44 (2.14-2.78)	2.30 (2.10-2.60)	4.40 (3.70-5.50)	9.70 (7.30-
20 years and older	03-04	2.41 (2.15-2.72)	2.60 (2.30-2.80)	5.10 (4.50-5.70)	9.50 (8.10-
	05-06	1.75 (1.62-1.89)	1.80 (1.70-2.00)	3.40 (3.10-3.70)	6.40 (5.80-
	07-08	1.99 (1.82-2.18)	2.00 (1.80-2.30)	3.90 (3.40-4.60)	7.40 (6.60-

CDC, Fourth National Exposure Report (2011)

EPA United States Environmental Protection Agency

Linking NHANES Urine Data and Exposure

Generalization of LaKind and Naiman (2008)



<u>*Steady-state assumption*</u>

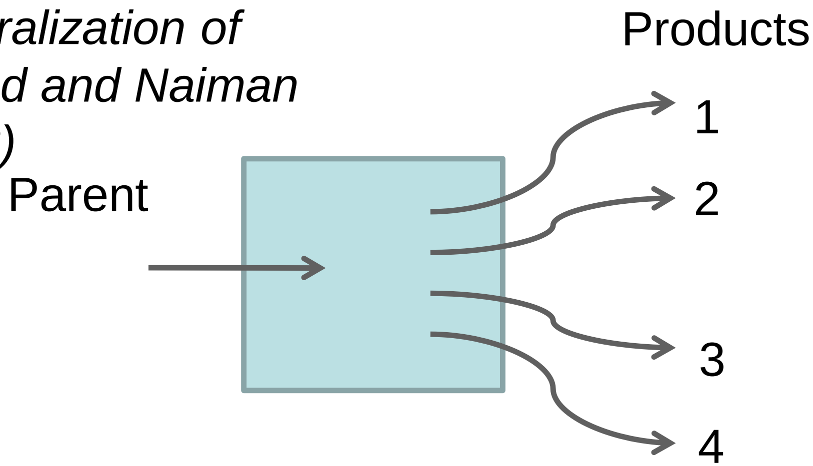
$$\left(\text{mg/kg/day}\right)_i = \frac{1}{70\ \text{kg}} \frac{\text{mg}_i}{\text{g}_{\text{creatine}}} * \frac{\text{g}_{\text{creatine}}}{\text{day}}$$

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

Method detailed in Wambaugh *et al.* (2013) Supplemental

Linking NHANES Urine Data and Exposure

Generalization of LaKind and Naiman (2008)



<u>*Steady-state assumption*</u>

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

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

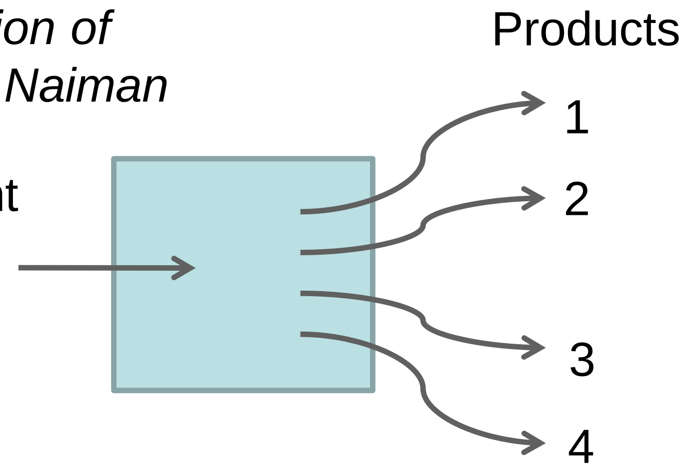
Observations (CDC NHANES urine samples)

Method detailed in Wambaugh *et al.* (2013) Supplemental

Linking NHANES Urine Data and Exposure

Generalization of LaKind and Naiman (2008)

Parent



Steady-state assumption

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

$$(\text{mg/kg/day})_0 = \text{MW}_0 \sum_i \phi_{0i} \frac{(\text{mg/kg/day})_i}{\text{MW}_i}$$

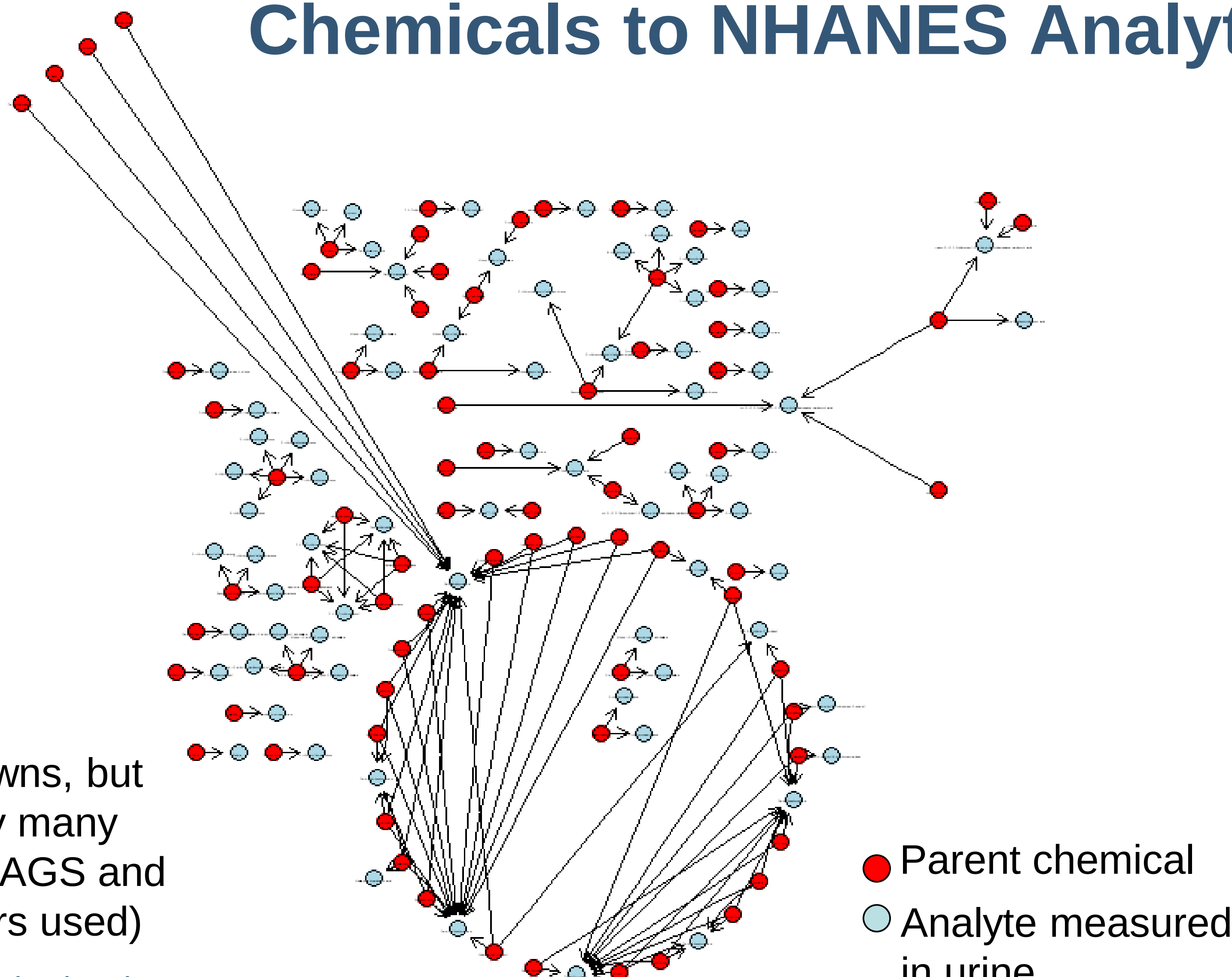
Observations (CDC NHANES urine samples)

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

Method detailed in Wambaugh *et al.* (2013) Supplemental

United States Environmental Protection Agency

Mapping Putative Parent Chemicals to NHANES Analytes

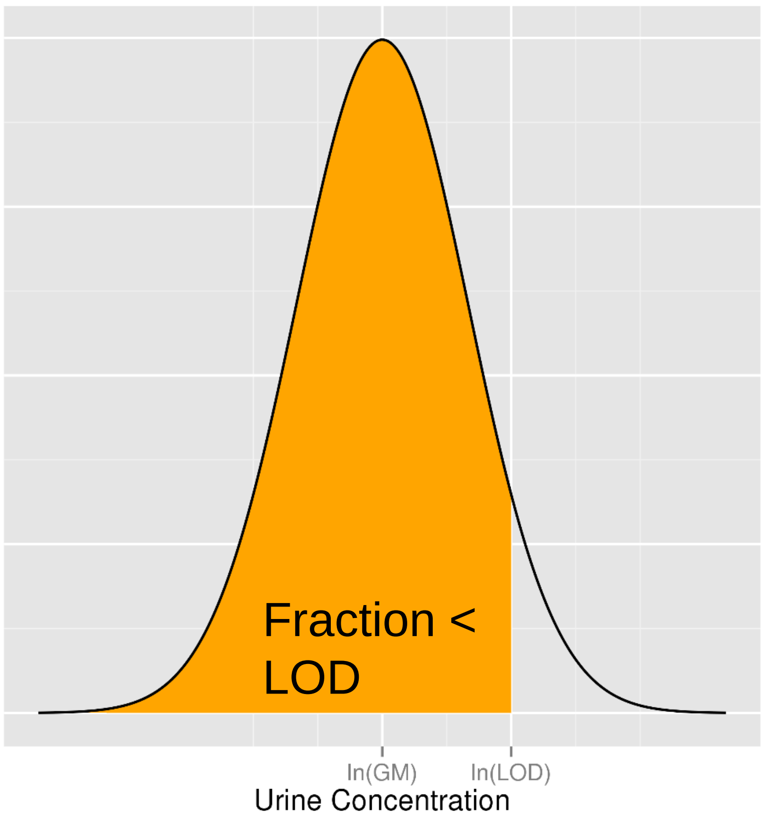


Lots of unknowns, but MCMC can try many possibilities (JAGS and STAN samplers used)

Method detailed in Wambaugh *et al.* (2013) Supplemental

Limit of Detection (LOD)

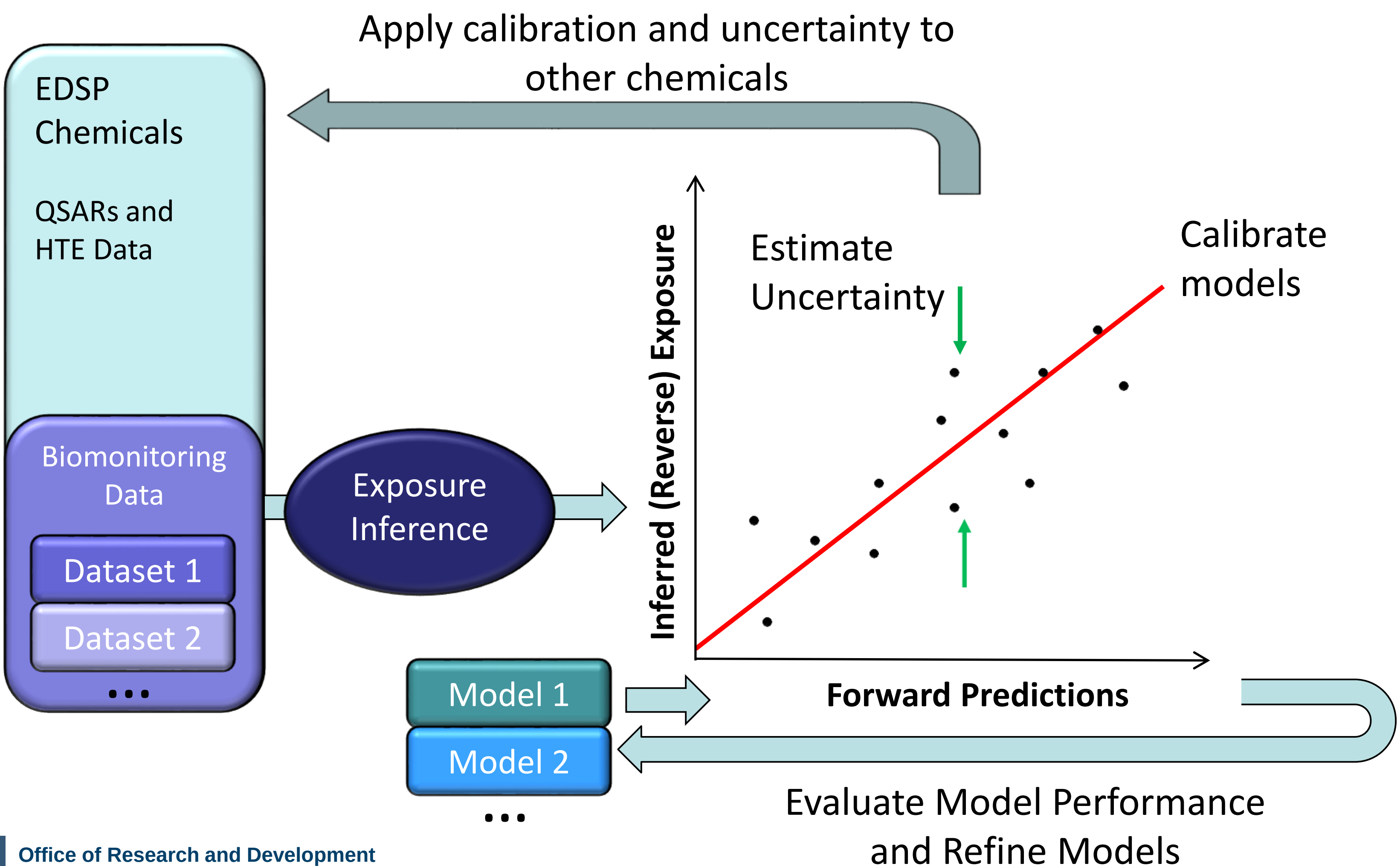
- If observations < analytic detection limits: We model the data as left censored observations from lognormal population distribution



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

Figure from Woodrow Setzer

Systematic Empirical Evaluation of Models



Statement of New Problem: Data Concerns

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

- What we would like to know is:

    - What are the few, most-easily obtained exposure heuristics that allow for prioritization?

Statement of New Problem: Data Concerns

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

- What we would like to know is:

  - What are the few, most-easily obtained exposure heuristics that allow for prioritization?

- What we can answer is this:

  - Given a variety of rapidly obtained data (putative use categories and physico-chemical properties, largely from QSAR) which data best explain exposure inferred from the available biomonitoring data?

  - Hoping to find simple heuristics for exposure *e.g.,* use in fragrances, use as a food additive, octanol:water partition coefficient, vapor pressure

Chemical Use Information for >30,000 Chemicals

ACToR UseDB: Chemical Use Categories estimated from ACToR (computational toxicology database):

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

Table: Hits per use category for a given chemical

CASRN	Category 1	Category 2	…	Category 12
65277-42-1	0	10	…	1
50-41-9	31	7	…	3
…	…	…	…	…

Binary matrix

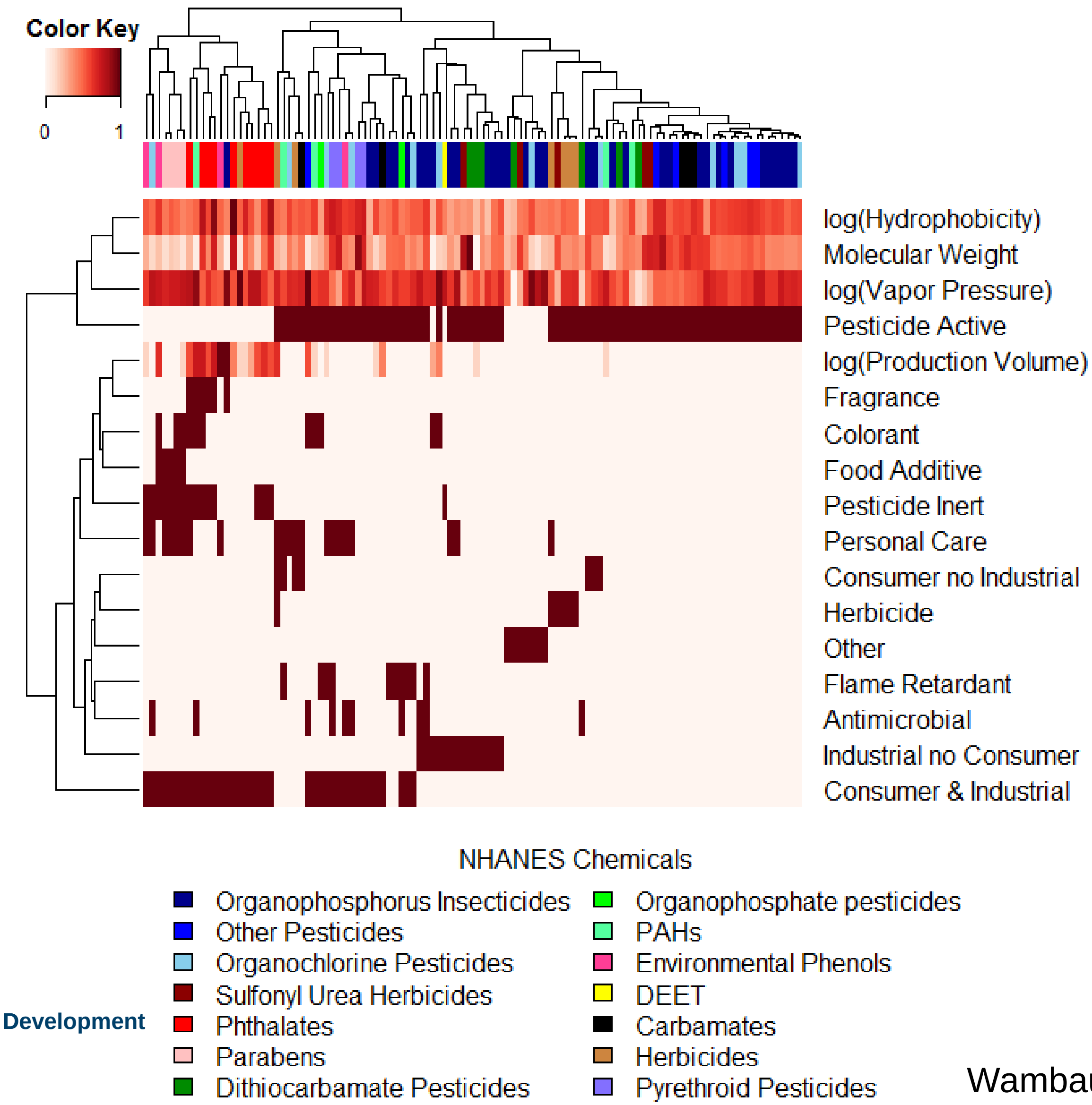
CASRN	Category 1	Category 2	…	Category 12
65277-42-1	0	1	…	0
50-41-9	1	1	…	0
…	…	…	…	…

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

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

Work by Alicia Frame, Kathie Dionisio, Richard Judson
Dionisio et al. *manuscript in preparation*

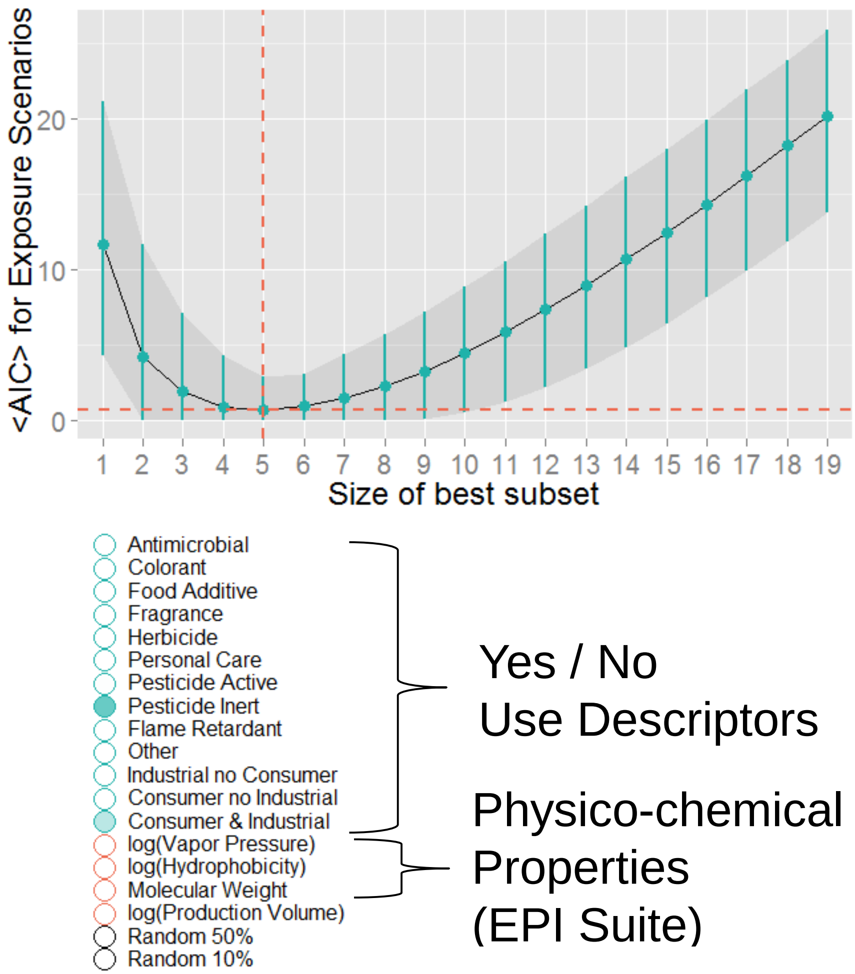
Heuristics for Chemical Use



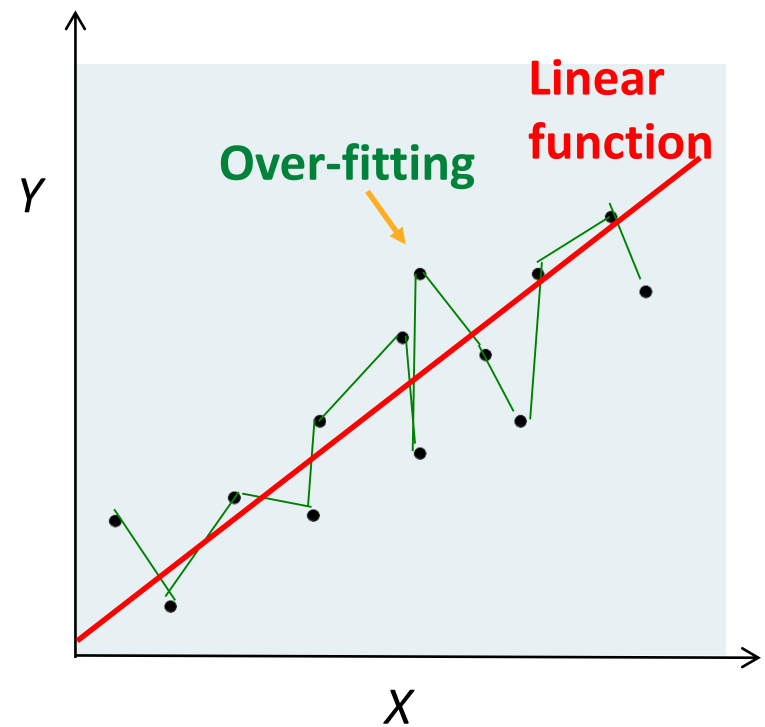
Wambaugh et al., (2014)

High Throughput Descriptors for Exposure

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

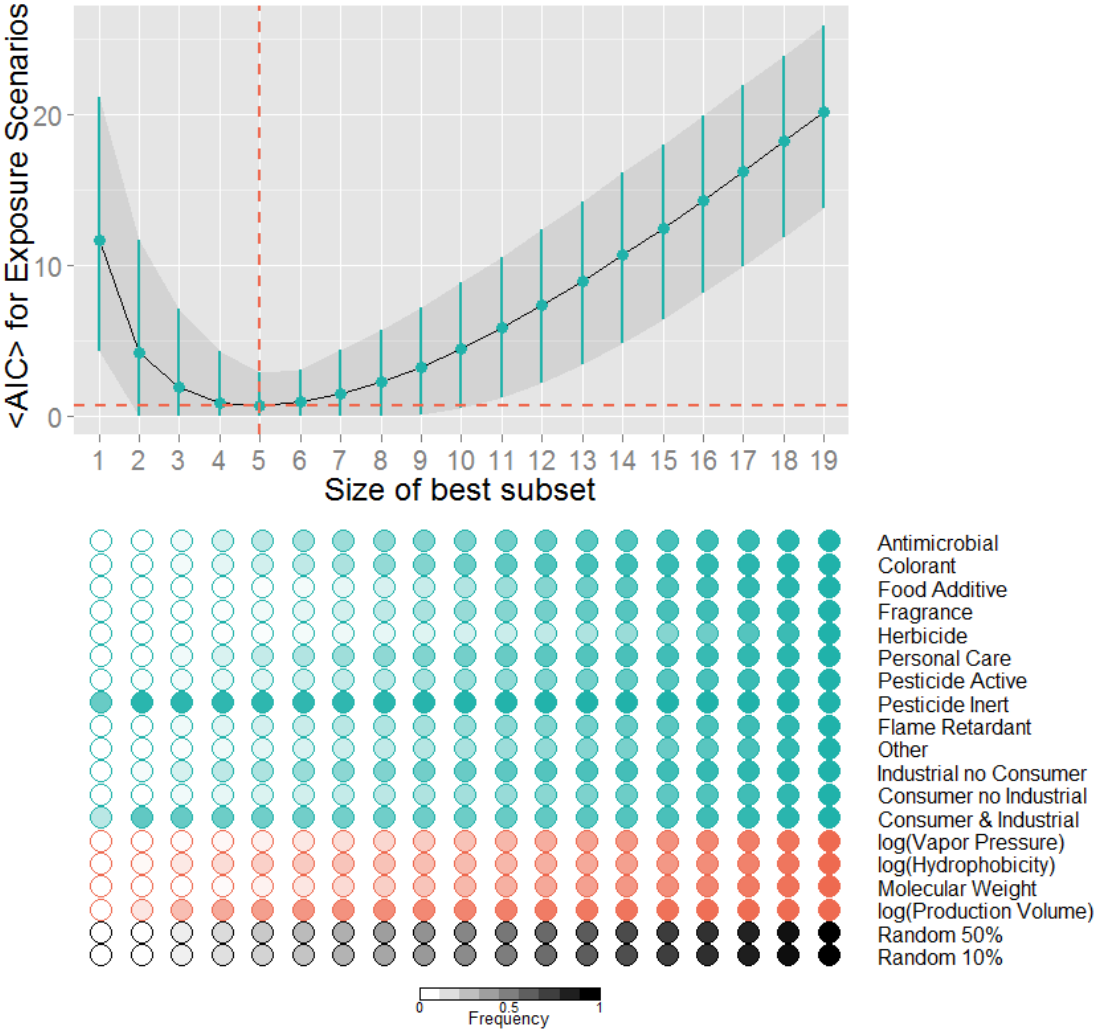


Noisy data and the danger of over-fitting



Wambaugh et al., (2014)

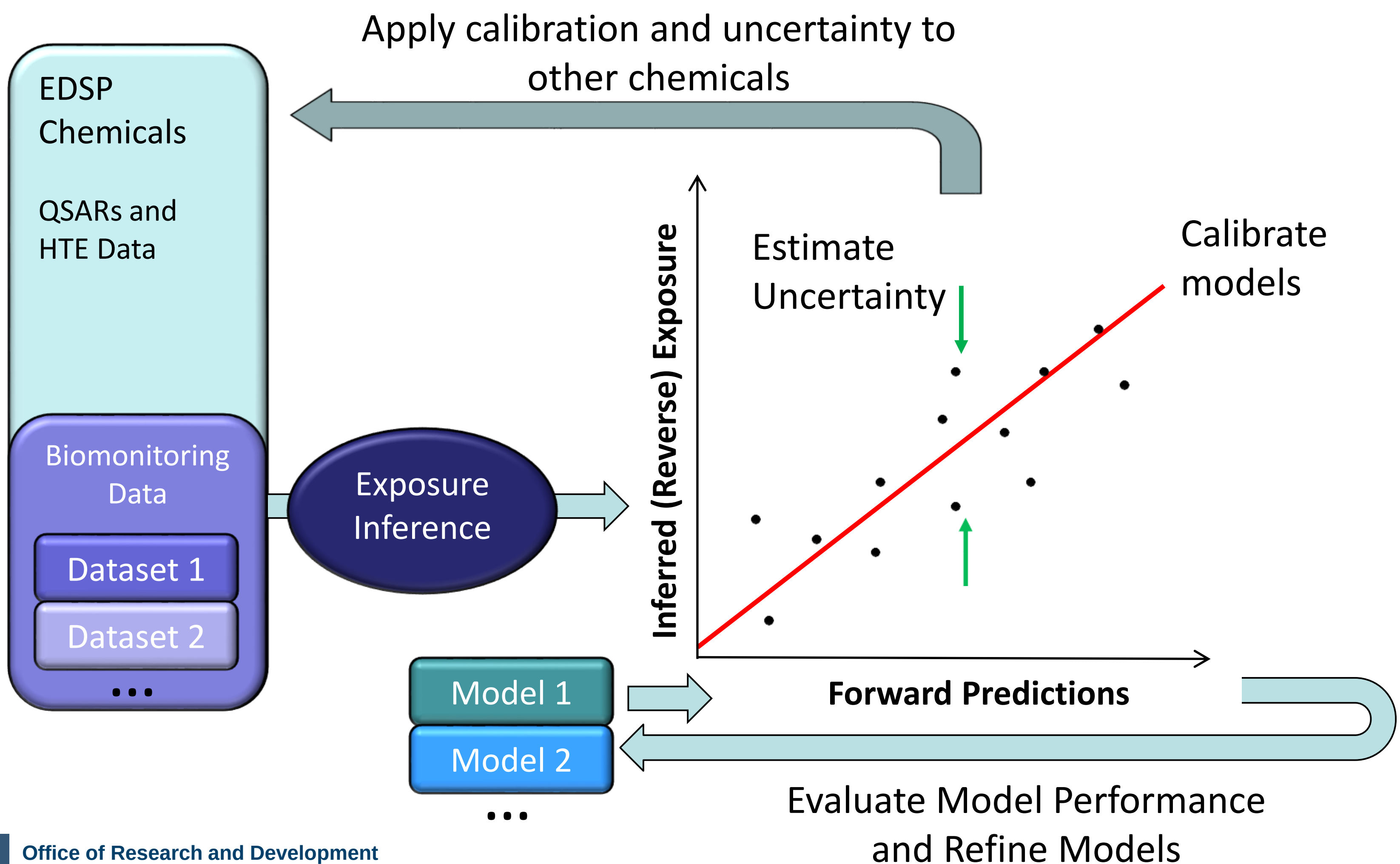
Not All Descriptors Are Useful



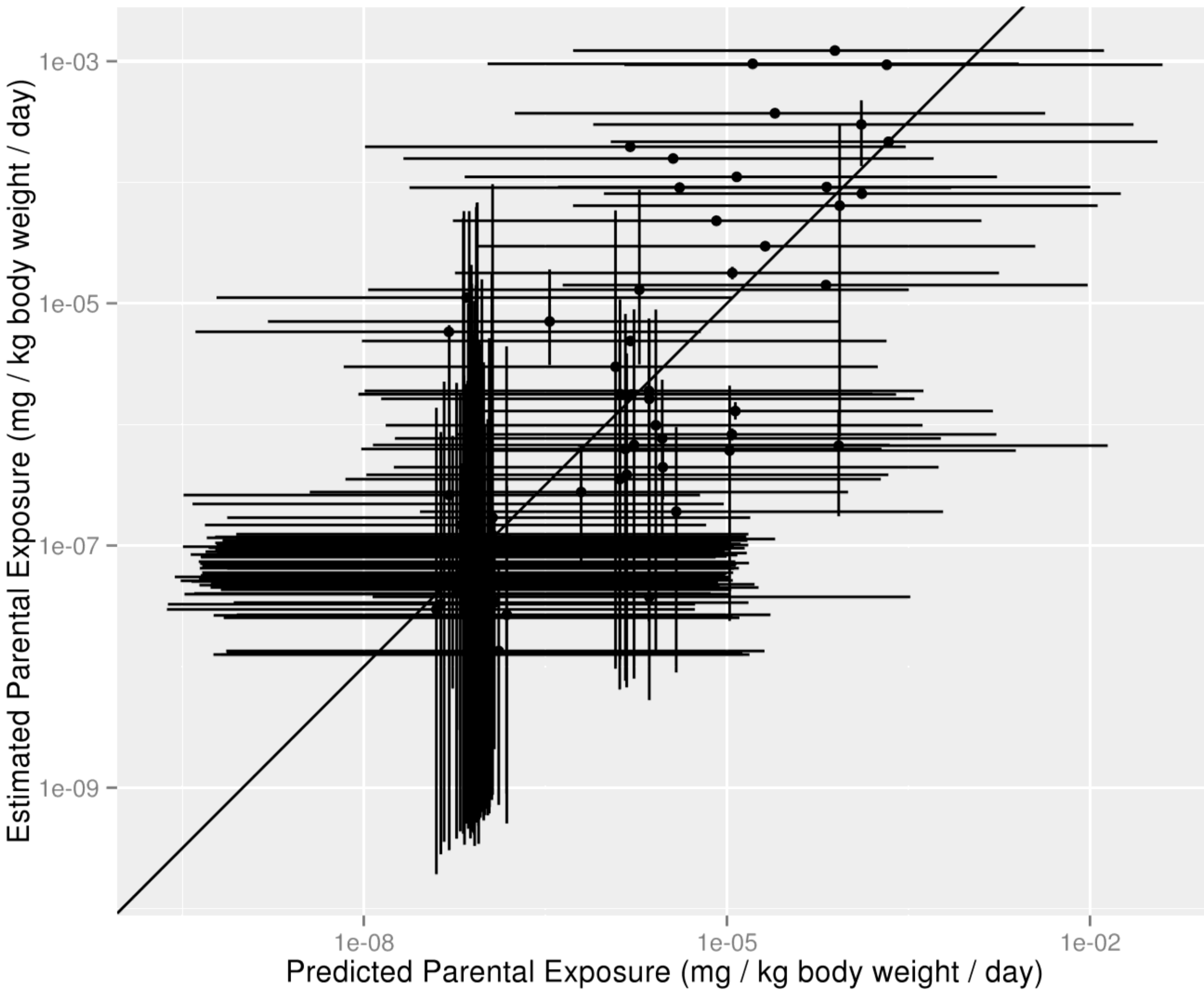
- The average relative AIC (smaller is better) for models made with different numbers of parameters for explaining 1500 different combinations of chemical exposures
- The predictors involved in the optimal model with higher frequencies are represented by darker circles, and those with lower frequencies by lighter circles
- As a sanity check, two random variables generated from binomial distribution with probability 50% and 10% of obtaining 1, are not selected as optimal descriptors in the five factor model

Wambaugh et al., (2014)

Systematic Empirical Evaluation of Models



Predicting NHANES exposure rates



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

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

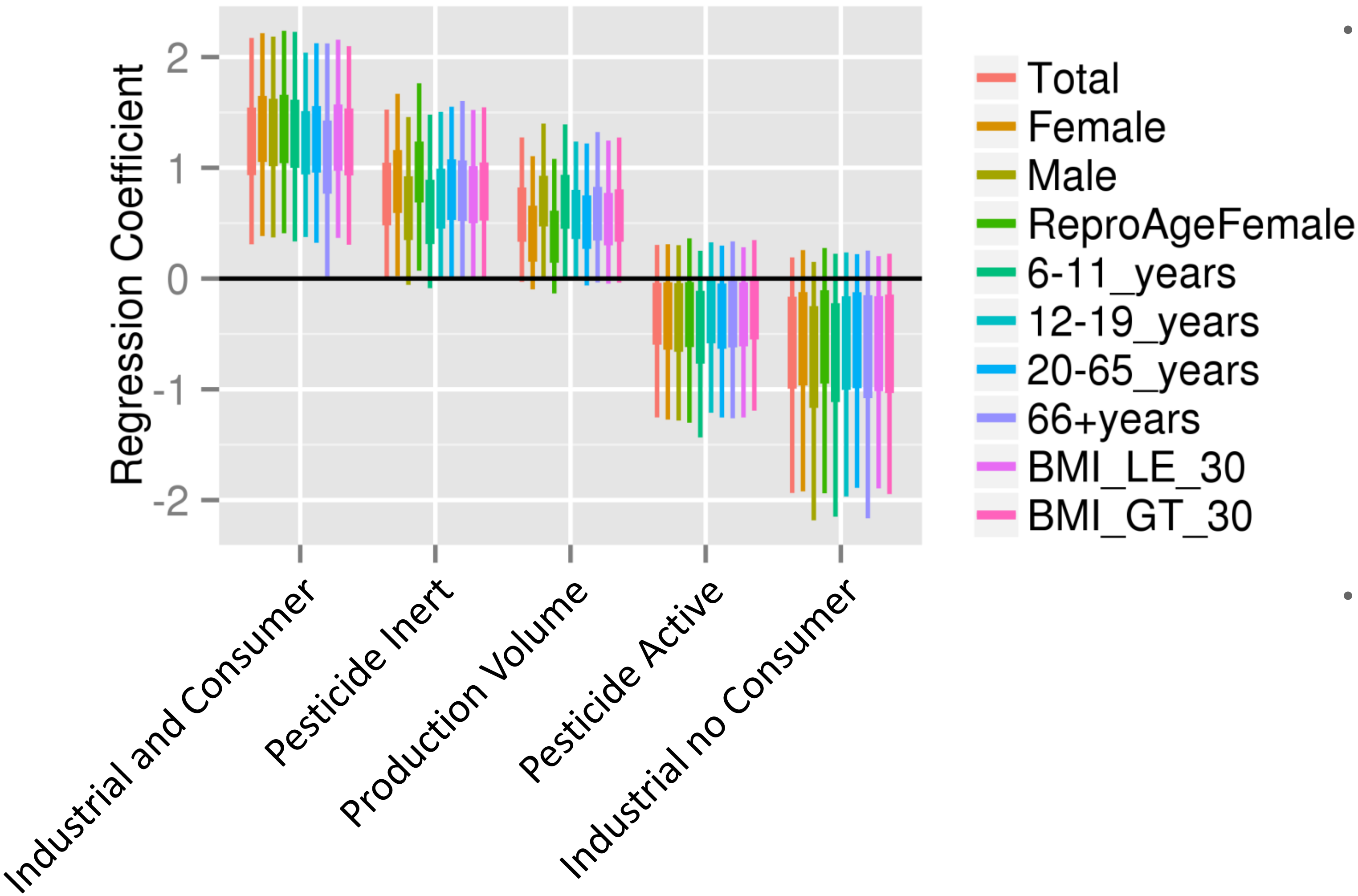
Wambaugh et al., (2014)

High-throughput exposure heuristics

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

Wambaugh et al., (2014)

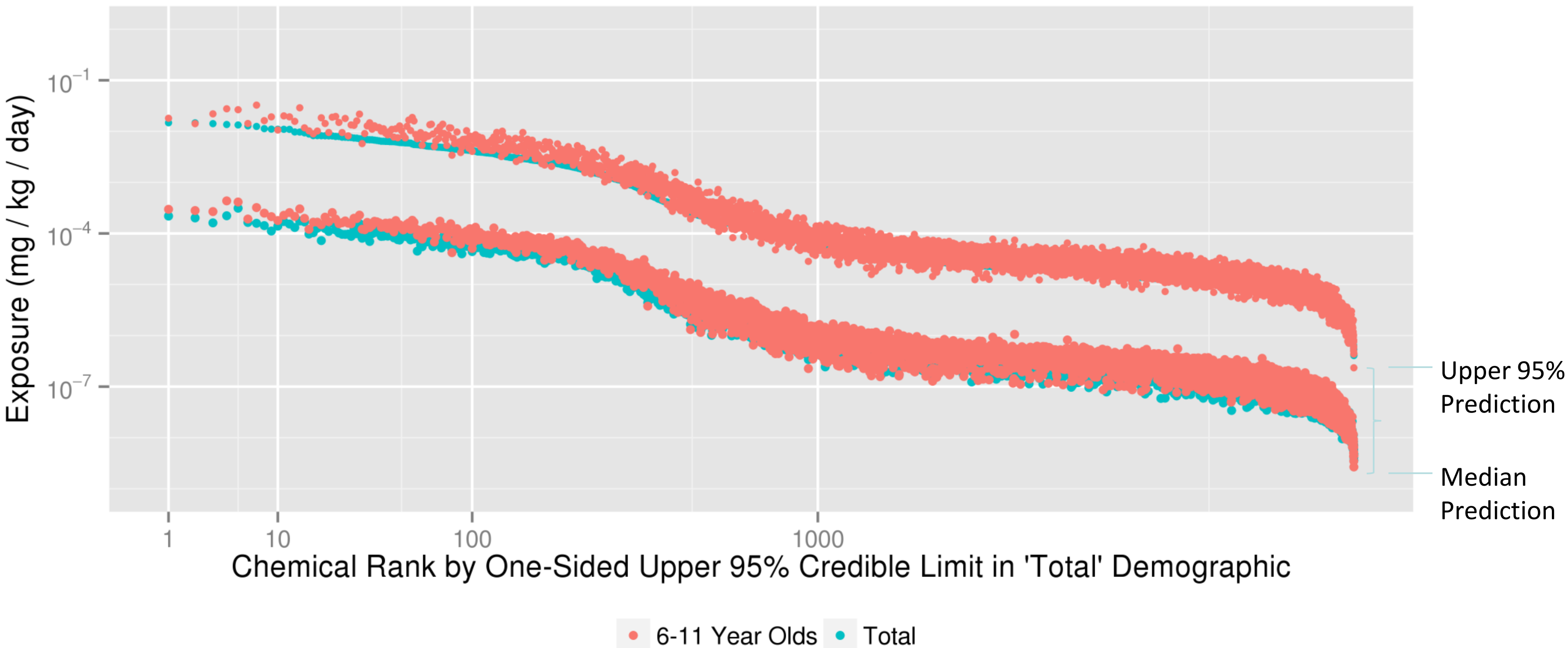
Predictors Do Not Vary Between Groups



- The vertical lines indicate the 95% credible interval across the 1500 different exposure scenarios inferred from the NHANES urine data
- SHEDS-HT (Isaacs et al., 2014) should help explain some remaining NHANES variability

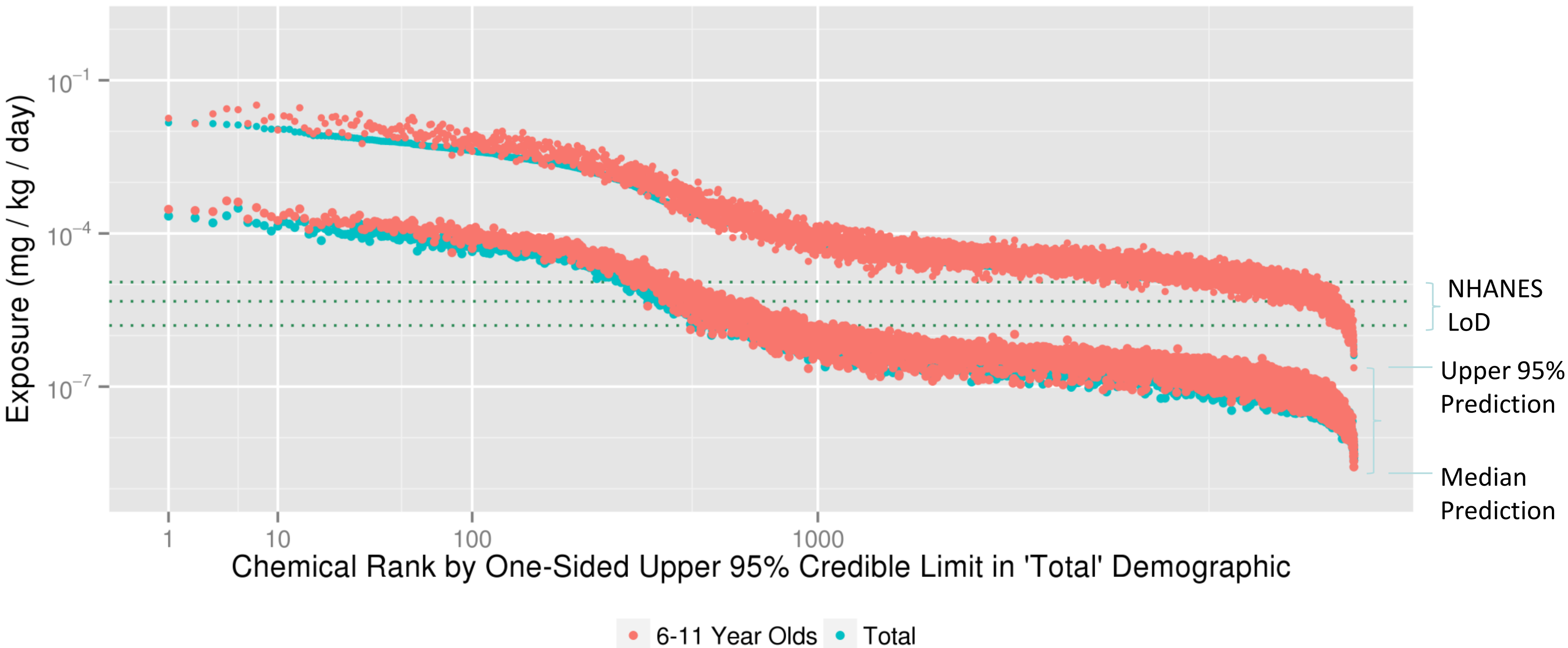
Wambaugh et al., (2014)

Calibrated Exposure Predictions for 7968 Chemicals



Wambaugh et al., (2014)

Calibrated Exposure Predictions for 7968 Chemicals

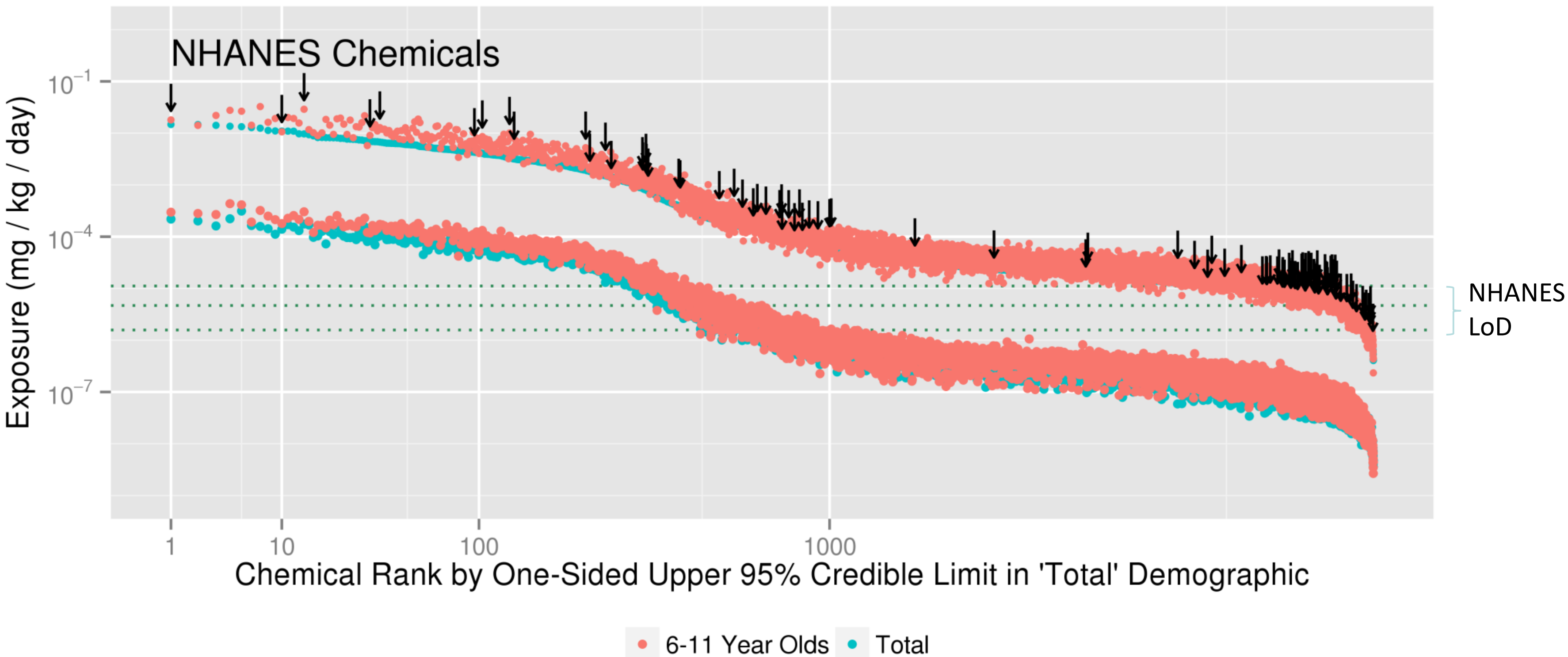


- We focus on the median and upper 95% predictions because the lower 95% is below the NHANES limits of detection (LoD)
- Dotted lines indicate 25%, median, and 75% of the LoD distribution

Wambaugh et al., (2014)

Calibrated Exposure Predictions for 7968 Chemicals

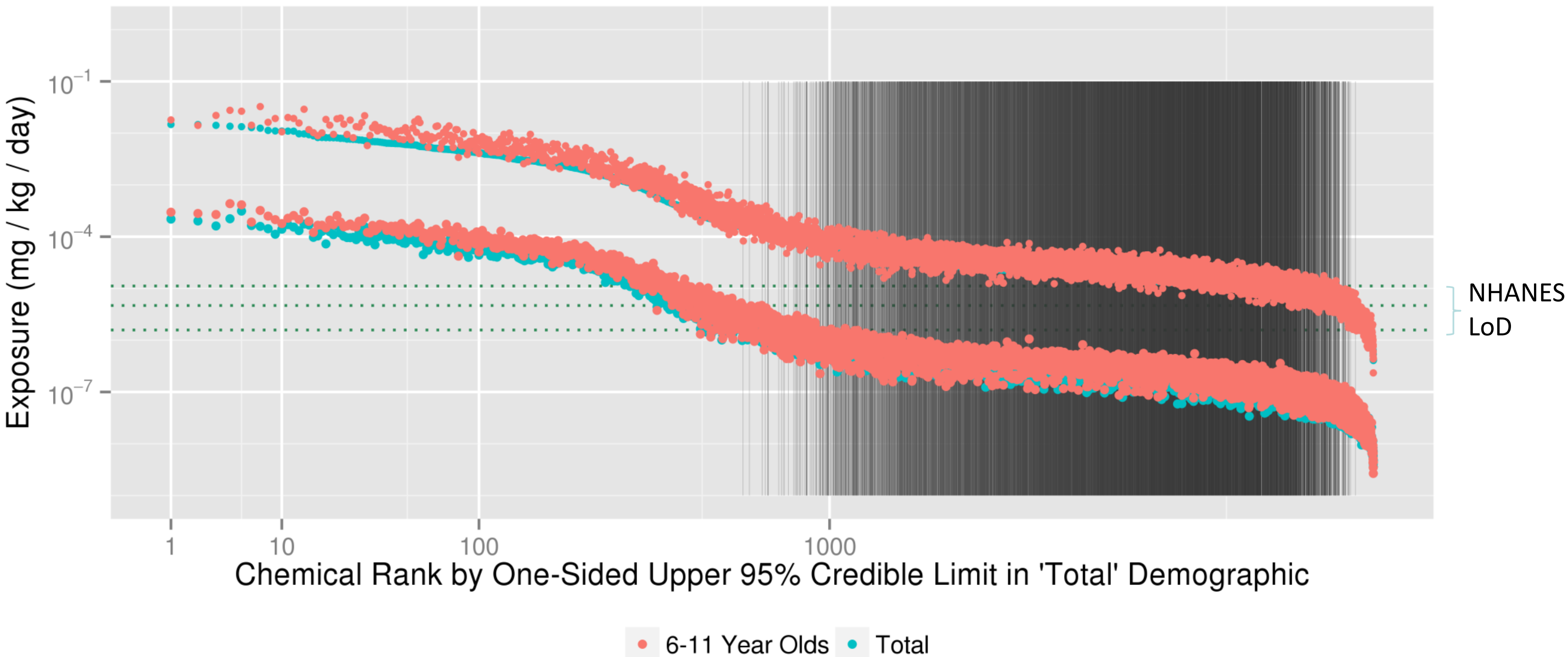
United States Environmental Protection Agency



- Chemicals currently monitored by NHANES are distributed throughput the predictions
- Chemicals with the first and ninth highest 95% limit are monitored by NHANES

Wambaugh et al., (2014)

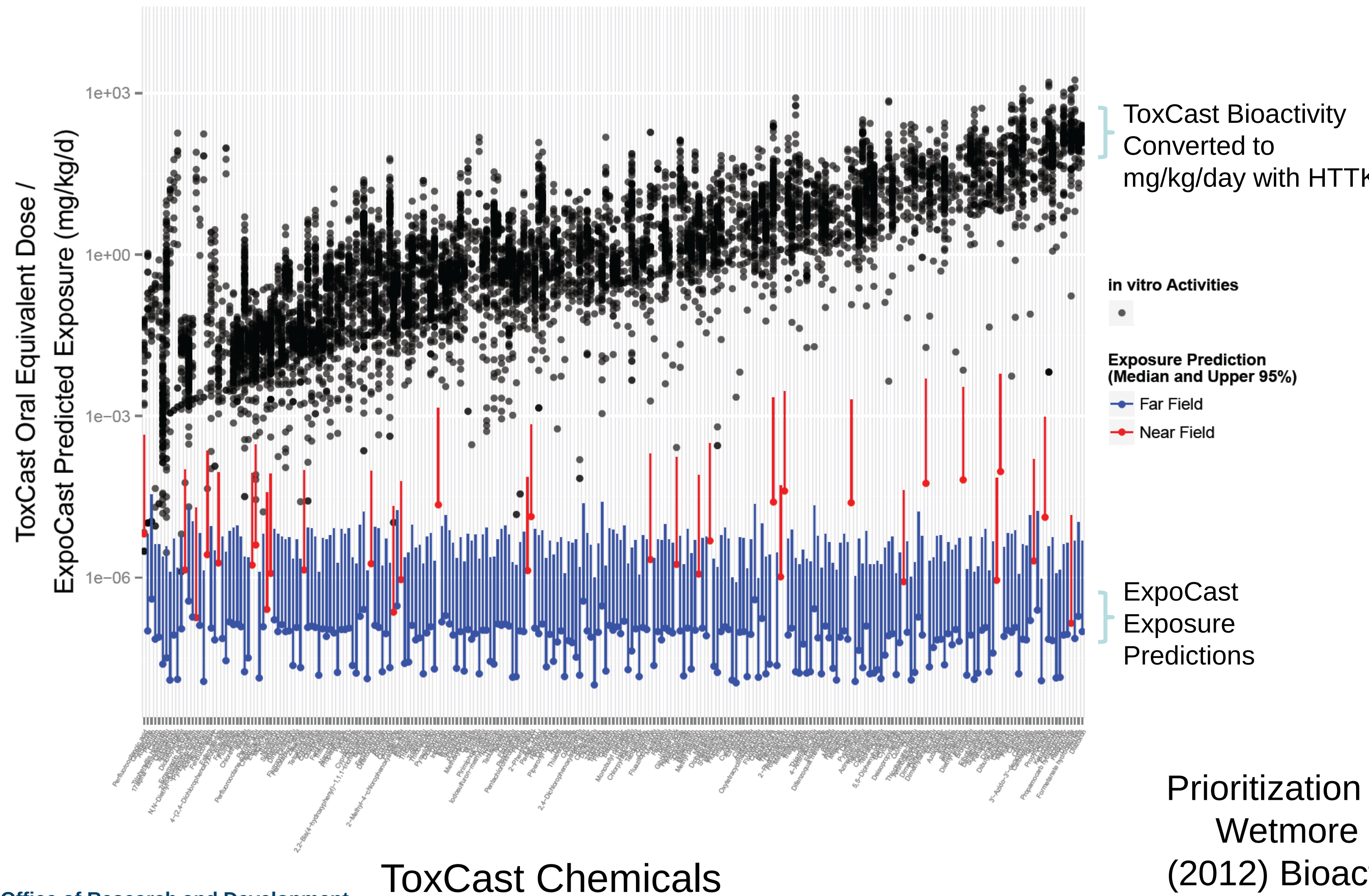
Calibrated Exposure Predictions for 7968 Chemicals



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

Wambaugh et al., (2014)

High Throughput Risk Prioritization



Prioritization as in Wetmore *et al*. (2012) Bioactivity, Dosimetry, and Exposure Paper

A Closer Look at Bisphenol A

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

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

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

*ExpoCast makes single prediction for Age 20-59y

A Closer Look at Triclosan

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

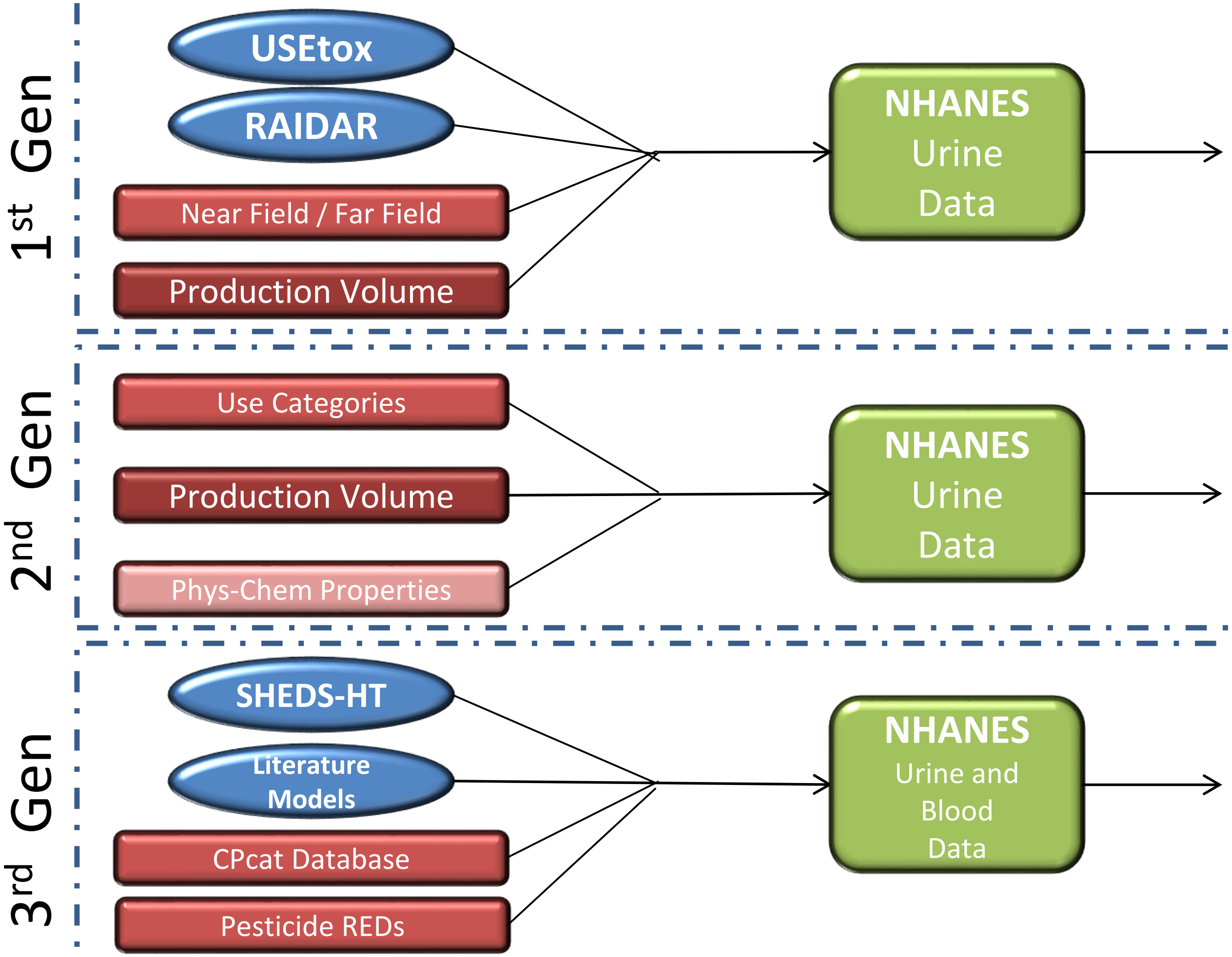
Demographic	Mage (2007)	Schafer (2004)	Geigy (1981) Mean	Geigy (1981) 95%	ExpoCast Geometric Mean Median	ExpoCast Geometric Mean Upper 95%
Total	2.5	2.9	2.9	4.5	0.0012	0.085
Age 6-11	1.6	1.9	1.7	2.4	0.0079	0.17
Age 12-19	2.7	3.2	4.1	6.2	0.0015	0.11
Age 20-59	2.9	3.2	3.0	4.7	0.0015	0.11
Age >=60	1.9	2.2	2.1	3.3	0.002	0.083
Male	3.1	3.8	3.6	5.6	0.0011	0.074
Female	2.0	2.1	2.1	3.4	0.0016	0.11

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

United States Environmental Protection Agency

SEEM Evolution

Model and Predictors | **Calibration/Evaluation Data** | **SEEM Conclusion**



1st Gen

- Existing complex fate and transport models have low correlation to measured exposures
- Near field factor most important

2nd Gen

- Simple, readily available data
- Better correlation to measured exposures
- Similar predictions across demographics

3rd Gen

- Isaacs et al. (2014) developed SHEDS-HT and predicted consumer product-driven exposures for 2507 chemicals
- SEEM analysis in progress

Better Models and Data Should Reduce Uncertainty

Uncertainty/Variability of NHANES Biomonitoring

~10% Far field (Industrial) Releases

~60% Indoor / Consumer Use

Consumer product database and two new near field models in 2014

Contents lists available at ScienceDirect

Food and Chemical Toxicology

journal homepage: www.elsevie

Development of a consumer product ingre
exposure screening and prioritization

M.-R. Goldsmith [a,*], C.M. Grulke [a], R.D. Brooks [b], T.R. T
R. Edwards [d], D.T. Chang [a], R. Tornero-Velez [a], K. Isaacs
J. Mitchell [g], D.A. Vallero [a], L. Phillips [a], M. Phillips [a], J.F
T.J. Buckley [a], C.C. Dary [a]

[a] United States Environmental Protection Agency (U.S. EPA), Office of Research and Develop
[b] Student Services Contractor at U.S. EPA, RTP, NC, United States
[c] Lockheed-Martin Information Technology, RTP, NC 27711, United States
[d] North Carolina State University, 2200 Hillsborough St., Raleigh, NC 27695, United States
[e] Oak Ridge Institute for Science and Education Fellow, United States
[f] University of North Carolina at Chapel Hill, Chapel Hill, NC 27514, United States
[g] Biosystems and Agricultural Engineering, Michigan State University, E. Lansing, MI 48824,

A R T I C L E I N F O

Article history:
Received 16 October 2013
Accepted 19 December 2013
Available online 27 December 2013

Keywords:
Chemical exposure
Consumer products
Ingredients
Product formulation
Near field exposure
Exposure prioritization

A B S T R A C T

Consumer products are a prin
able on the chemical ingredie
ent. To address this data gap
Material Safety Data Sheets (
sents 1797 unique chemicals
uct "use categories" within a
discuss ways in which it will
formulations for several indo
selection for monitoring near
uitous exposure sources usin
and across multiple consume
fied. Our database is publicly
predictive screening of chem
risk.

Environmental Science & Technology

Article

pubs.acs.org/est

SHEDS-HT: An Integrated Probabilistic Exposure Model for Prioritizing Exposures to Chemicals with Near-Field and Dietary Sources

Kristin K. Isaacs,[*,†] W. Graham Glen,[‡] Peter Egegh
Daniel Vallero,[†] Raina Brooks,[∥] Christopher M. Gr

[†] U.S. Environmental Protection Agency, Office of Research and
Alexander Drive, Research Triangle Park, North Carolina 27709,
[‡] Alion Science and Technology, 1000 Park Forty Plaza Suite 20
[§] Chemical Computing Group, Suite 910, 1010 Sherbrooke Stree
[∥] Student Services Contractor at U.S. Environmental Protection A
Carolina 27709, United States
[⊥] Lockheed Martin, 109 T.W. Alexander Drive, Research Triangle

Supporting Information

ABSTRACT: United States Environmental Protection Agenc
(USEPA) researchers are developing a strategy for high
throughput (HT) exposure-based prioritization of chemica
under the ExpoCast program. These novel modeling approache
for evaluating chemicals based on their potential for biologicall
relevant human exposures will inform toxicity testing an
prioritization for chemical risk assessment. Based on probabilisti
methods and algorithms developed for The Stochastic Huma
Exposure and Dose Simulation Model for Multimedia, Mult
pathway Chemicals (SHEDS-MM), a new mechanistic modelin

Environmental Science & Technology

Article

pubs.acs.org/est

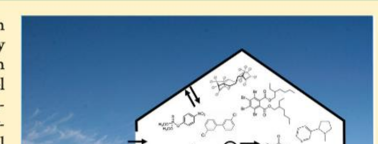
Model for Screening-Level Assessment of Near-Field Human Exposure to Neutral Organic Chemicals Released Indoors

Xianming Zhang,[*,†,§] Jon A. Arnot,[*,†,‡] and Frank Wania[†]

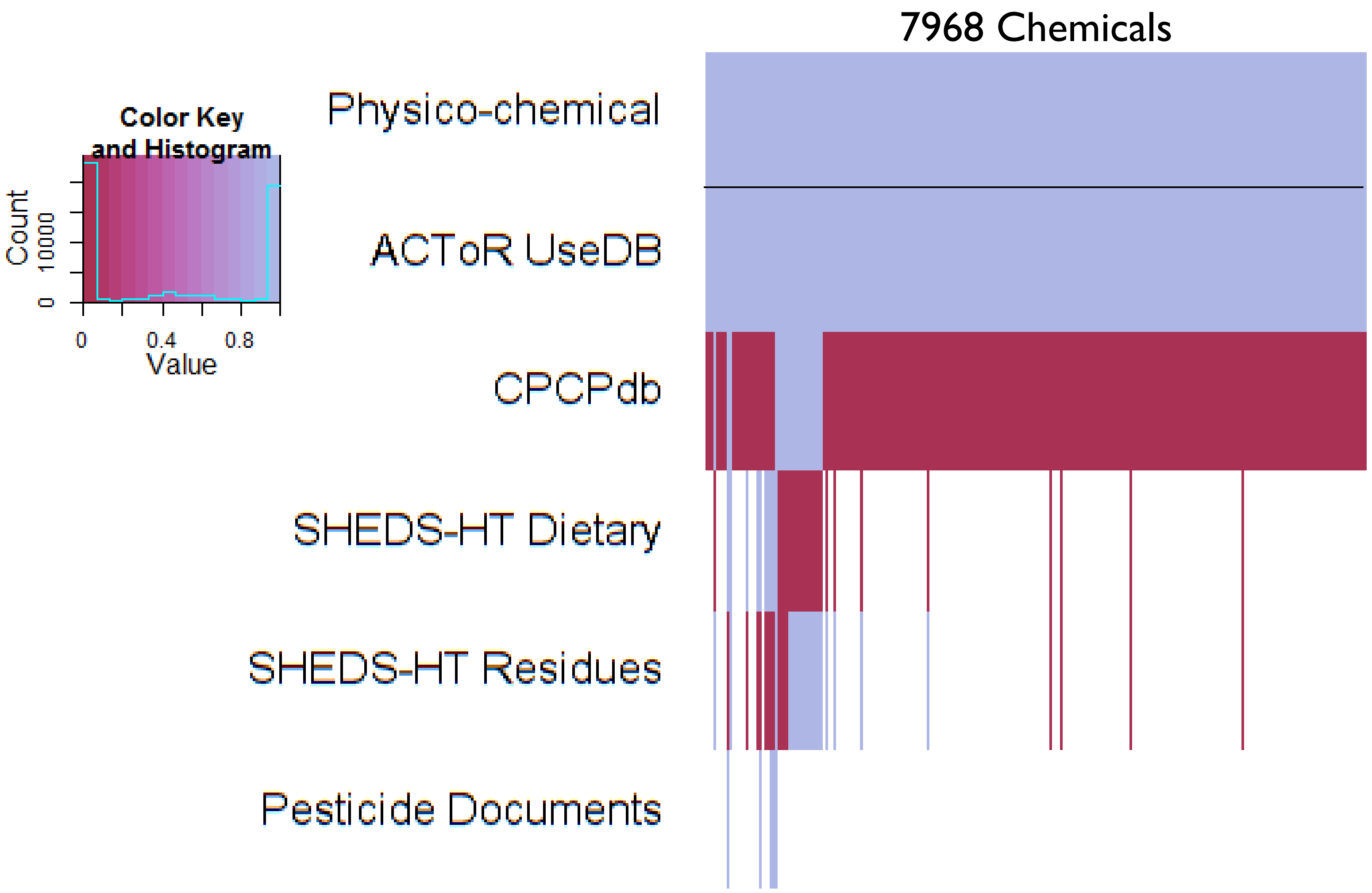
[†] Department of Physical and Environmental Sciences, University of Toronto Scarborough, Toronto, Ontario M1C 1A4, Canada
[‡] ARC Arnot Research and Consulting, Toronto, Ontario M4M 1W4, Canada

Supporting Information

ABSTRACT: Screening organic chemicals for hazard and risk to human health requires near-field human exposure models that can be readily parametrized with available data. The integration of a model of human exposure, uptake, and bioaccumulation into an indoor mass balance model provides a quantitative framework linking emissions in indoor environments with human intake rates (iRs), intake fractions (iFs) and steady-



Data Inhomogeneity



Conclusions

United States Environmental Protection Agency

- We identify those HTE factors that correlate with the NHANES data and estimate uncertainty
- The calibrated meta-model can estimate relative levels of chemical exposures for 7968 chemicals
  - This includes thousands of chemicals with no other data on human exposure
  - Same factors are predictive (R^2 ~ 0.5) across demographics characterized by NHANES
- Different demographics have different mean (overall) exposures:
  - There are demographic-specific aspects not currently described by available HTE factors
- Upcoming analysis:
  - Augment heuristics with calibrations of new mechanistic HT models for exposure from consumer use and indoor environment (*e.g.*, SHEDS-HT)
  - Develop new data sources with additional chemical descriptors (*e.g.*, CPcatDB)
  - Should help decrease uncertainties and increase confidence in extrapolation

Acknowledgements

EPA Office of Research and Development

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

NRMRL
Jane Bare
Xiaoyu Liu

NHEERL
Jane Ellen Simmons
Marina Evans
Mike Hughes

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

*Trainees

Arnot Research and Consulting
Jon Arnot

Chemical Computing Group
Rocky Goldsmith

Environmental Protection Agency
Alicia Frame

Hamner Institutes
Barbara Wetmore
Cory Strope

Indiana University
James Sluka

Michigan State University
Jade Mitchell

National Institute for Environmental Health Sciences (NIEHS)
Nisha Sipes
Kyla Taylor
Kristina Thayer

Netherlands Organisation for Applied Scientific Research (TNO)
Sieto Bosgra

North Carolina State University
Anran Wang

University of California, Davis
Deborah Bennett

University of Michigan
Olivier Jolliet

University of North Carolina, Chapel Hill
Alexander Sedykh*
Alex Tropsha

United States Environmental Protection Agency

References

- Cohen Hubal, Elaine A., et al. "Advancing exposure characterization for chemical evaluation and risk assessment." *Journal of Toxicology and Environmental Health, Part B* 13.2-4 (2010): 299-313.
- Dionisio, Kathie L., et al. “Exploring Consumer Exposure Pathways and Patterns of Use for Chemicals in the Environment.” *submitted.*
- Egeghy, Peter P., et al. "The exposure data landscape for manufactured chemicals." *Science of the Total Environment* 414 (2012): 159-166.
- 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.
- Isaacs, Kristin K., et al. "SHEDS-HT: An Integrated Probabilistic Exposure Model for Prioritizing Exposures to Chemicals with Near-Field and Dietary Sources." *Environmental science & technology* 48.21 (2014): 12750-12759.
- LaKind, Judy S., and Daniel Q. Naiman. "Daily intake of bisphenol A and potential sources of exposure: 2005–2006 National Health and Nutrition Examination Survey." *Journal of Exposure Science and Environmental Epidemiology* 21.3 (2010): 272-279.
- Pilkey, Orrin H., and Linda Pilkey-Jarvis. *Useless arithmetic: why environmental scientists can't predict the future*. Columbia University Press, 2007.
- Schmidt, Charles W. "TOX 21: new dimensions of toxicity testing." *Environ Health Perspect* 117.8 (2009): A348-A353.
- Silver, Nate. *The signal and the noise: Why so many predictions fail-but some don't*. Penguin, 2012.
- Sobus, Jon R., et al. "A biomonitoring framework to support exposure and risk assessments." *Science of the Total Environment* 409.22 (2011): 4875-4884.
- Tan, Yu-Mei, et al. "Reconstructing human exposures using biomarkers and other “clues”." *Journal of Toxicology and Environmental Health, Part B* 15.1 (2012): 22-38.
- Wetmore, Barbara A., et al. "Integration of Dosimetry, Exposure, and High-Throughput Screening Data in Chemical Toxicity Assessment." *Toxicological Sciences* 125.1 (2012): 157-174.
- 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).