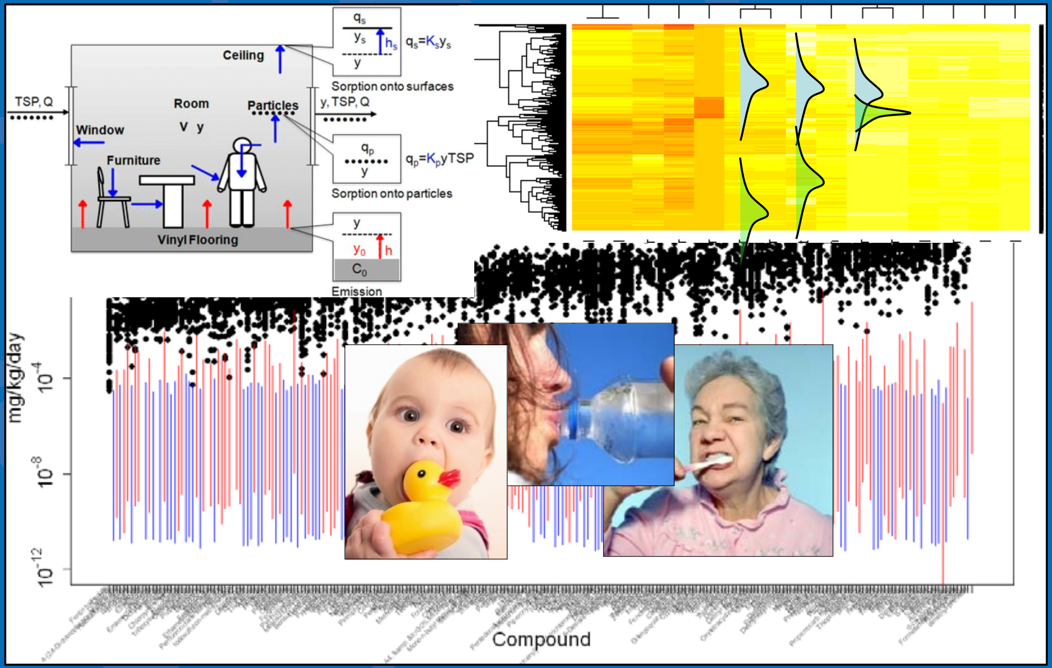
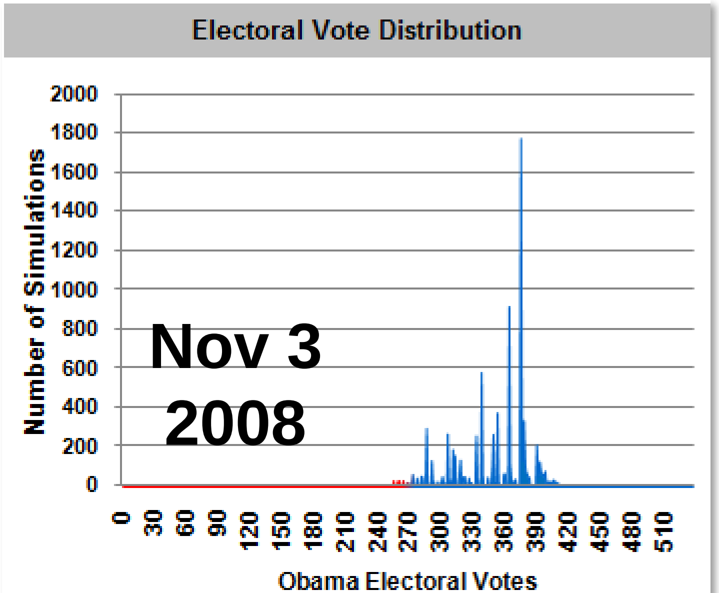
High Throughput Heuristics Can Forecast Human Exposure to Environmental Chemicals

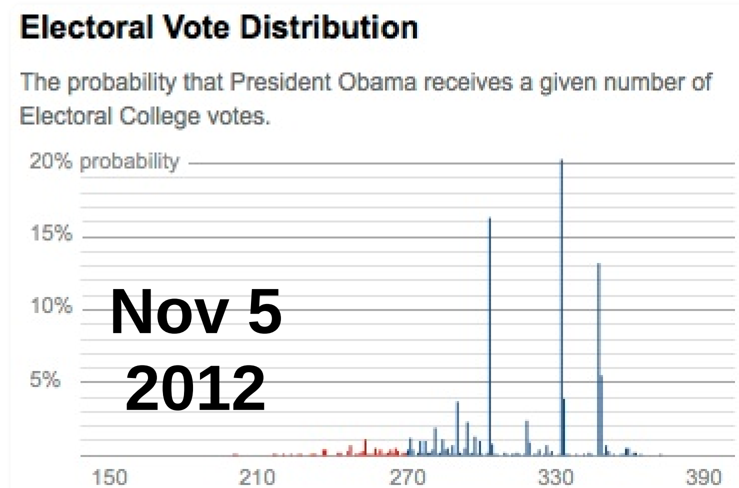
John Wambaugh
U.S. EPA, National Center for Computational Toxicology



June 27, 2013

The Signal and the Noise (2012)





Nate Silver (fivethirtyeight blog) has called the last two presidential elections correctly (a coin would do this one in four times)

He has called 99/100 state results correctly (a coin would do this one in $\sim 10^{28}$ times)

Nate Silver: *How to Make Good Forecasts*

1) Think probabilistically

2) Forecasts change – today's forecast reflects the best available data today

3) Look for consensus – multiple models/predictions

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"

Exposure Forecasting: ExpoCast

United States Environmental Protection Agency

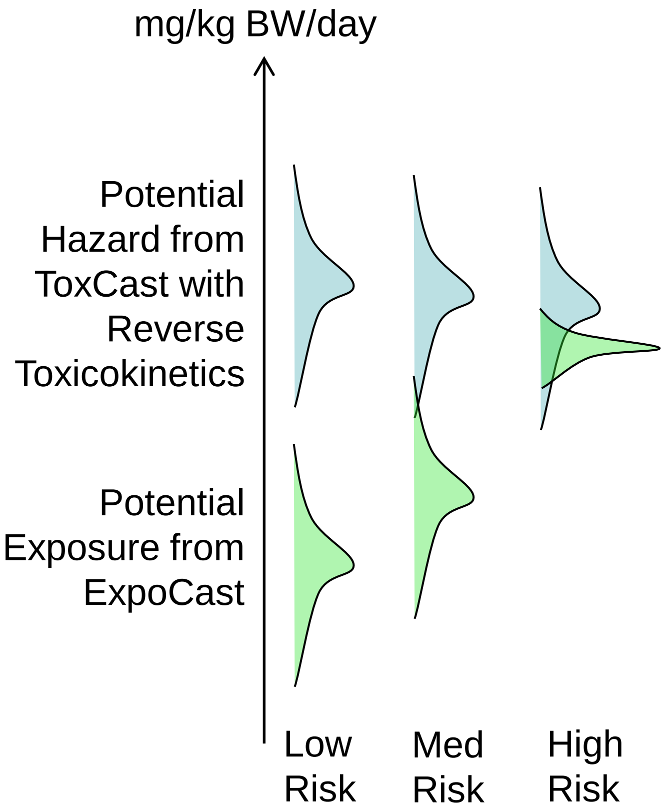
There are thousands of chemicals in commerce, most without enough data for risk evaluation

Risk is the product of hazard and exposure

High throughput *in vitro* methods beginning to bear fruit on potential hazard for many of these chemicals

Methods exist for approximately converting these *in vitro* results to daily doses needed to produce similar levels in a human (IVIVE)

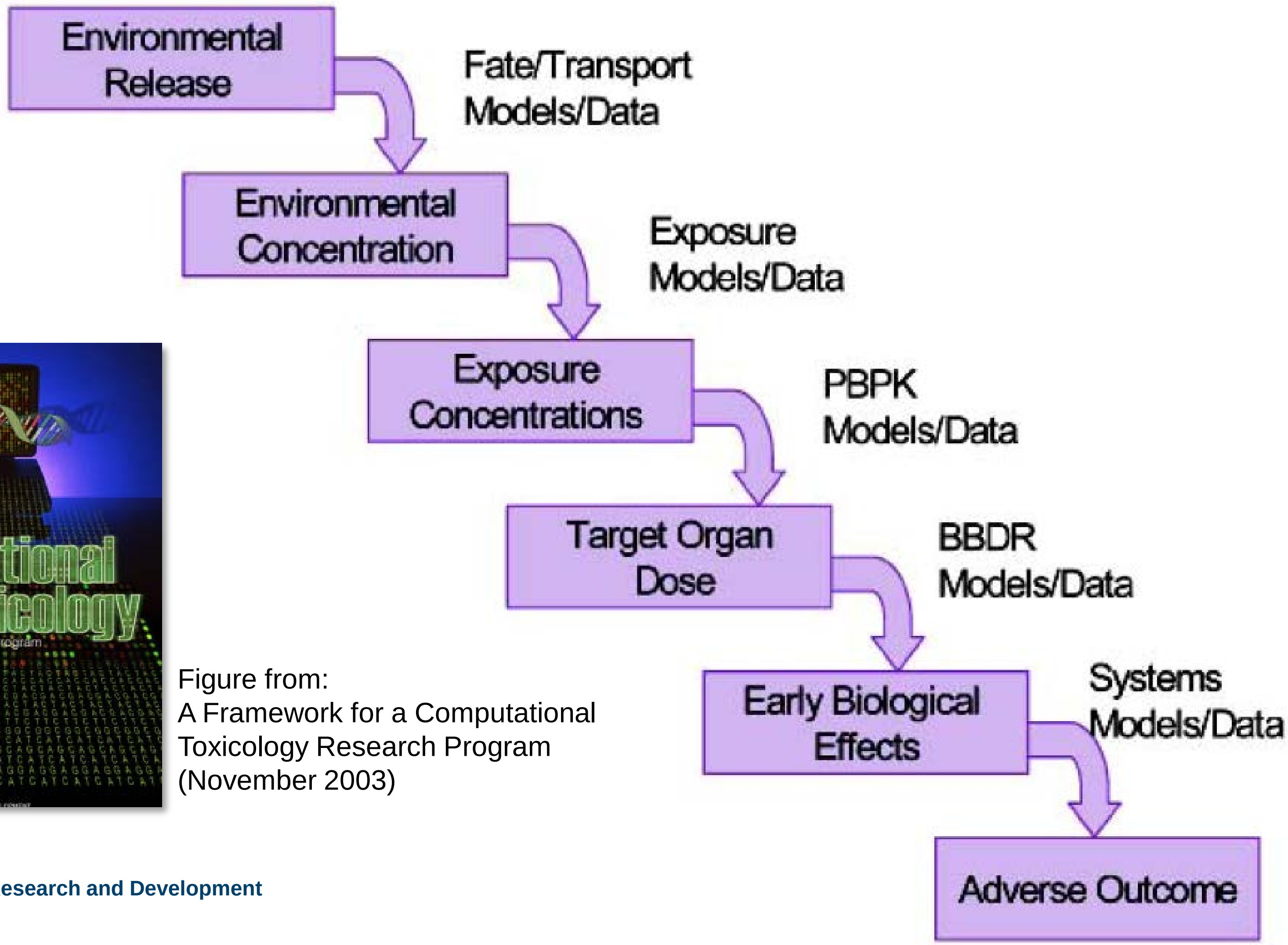
What can we say about exposure with the limited data we have?



e.g. Judson *et al.*, (2011) Chemical Research in Toxicology

What can we forecast about a new chemical based upon previously studied chemicals?

Source-to-Outcome Continuum



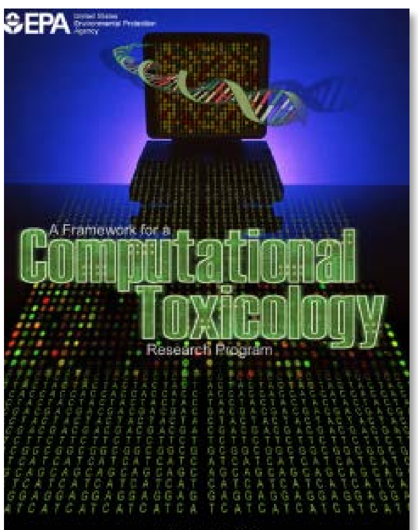
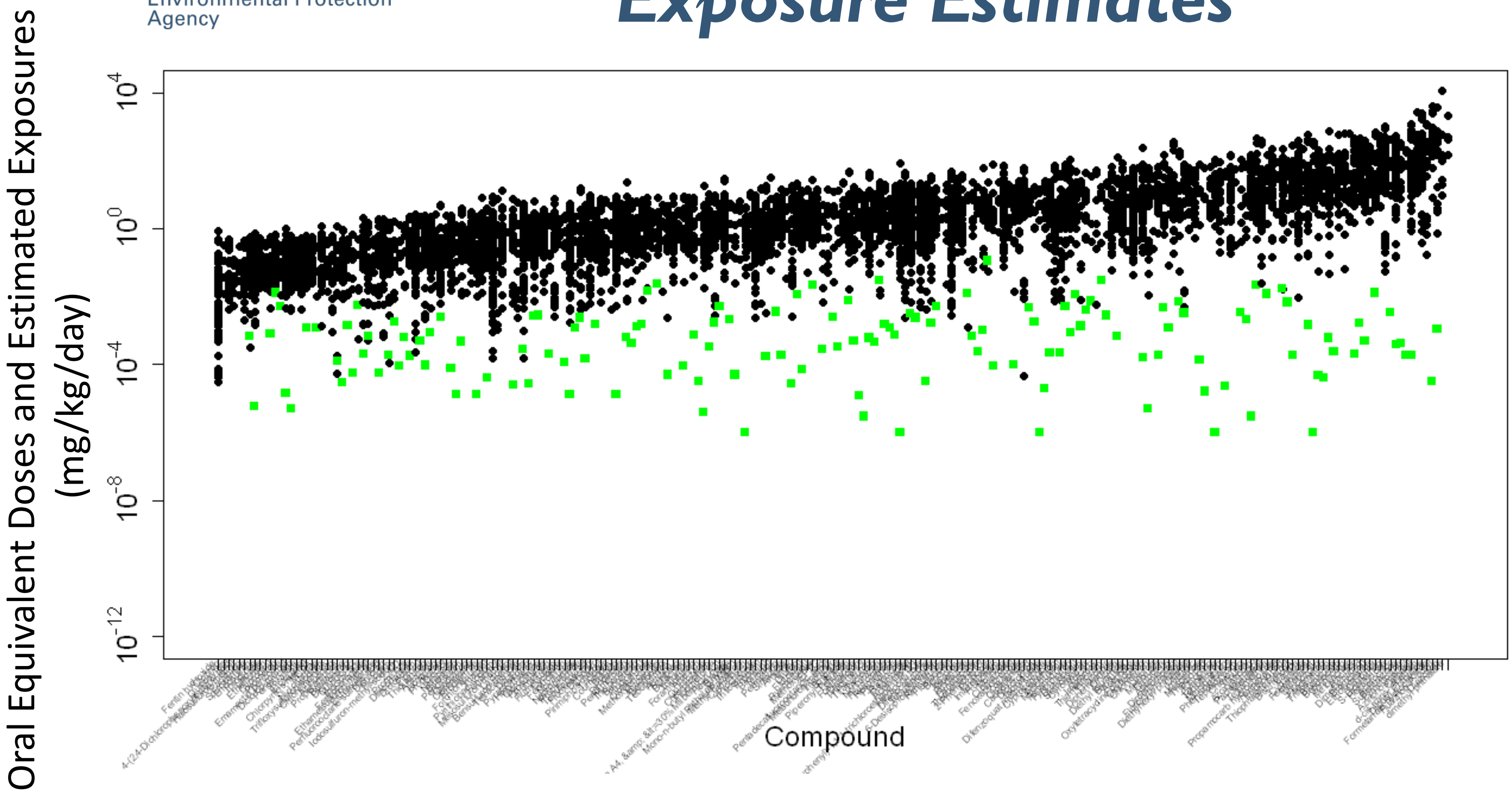


Figure from:
A Framework for a Computational Toxicology Research Program (November 2003)

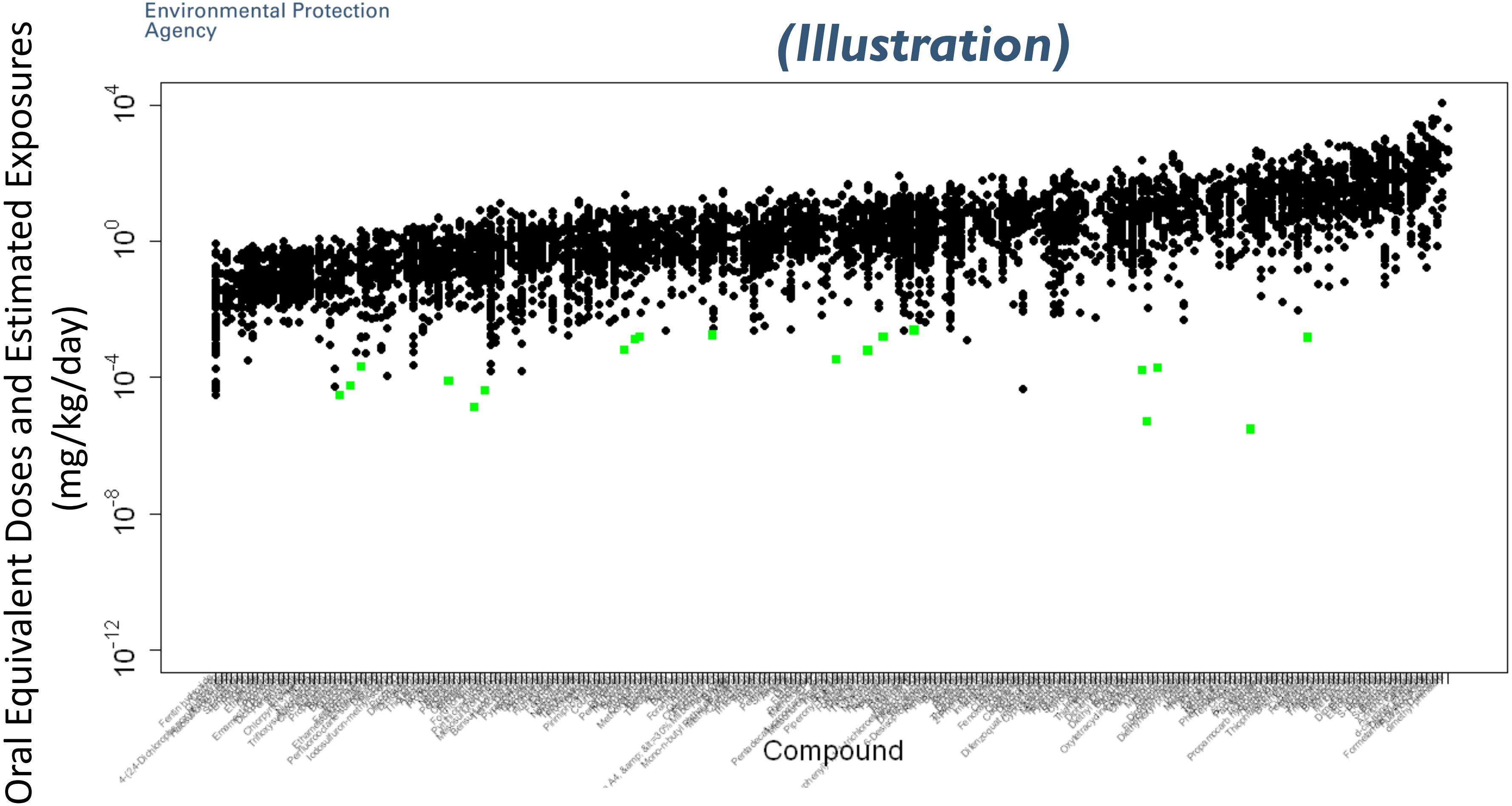
ToxCast Oral Equivalent Doses and Exposure Estimates



Green squares indicate estimated exposures from EPA REDs or CDC NHANES: ~71% of Phase I

Wetmore et al. Tox. Sci (2012)

The Exposure Coverage of the ToxCast Phase II Chemicals (Illustration)



Green squares indicate estimated exposures from EPA REDs or CDC NHANES: ~71% of Phase I
~7% of Phase II

United States Environmental Protection Agency

High Throughput Exposure Predictions

Goal: A high-throughput exposure approach to use with the ToxCast chemical hazard identification.

Proof of Concept: Using off-the-shelf models capable of quantitatively predicting exposure determinants in a high throughput (1000s of chemicals) manner and then evaluate those predictions to characterize uncertainty (Wambaugh *et al.*, ES&T)

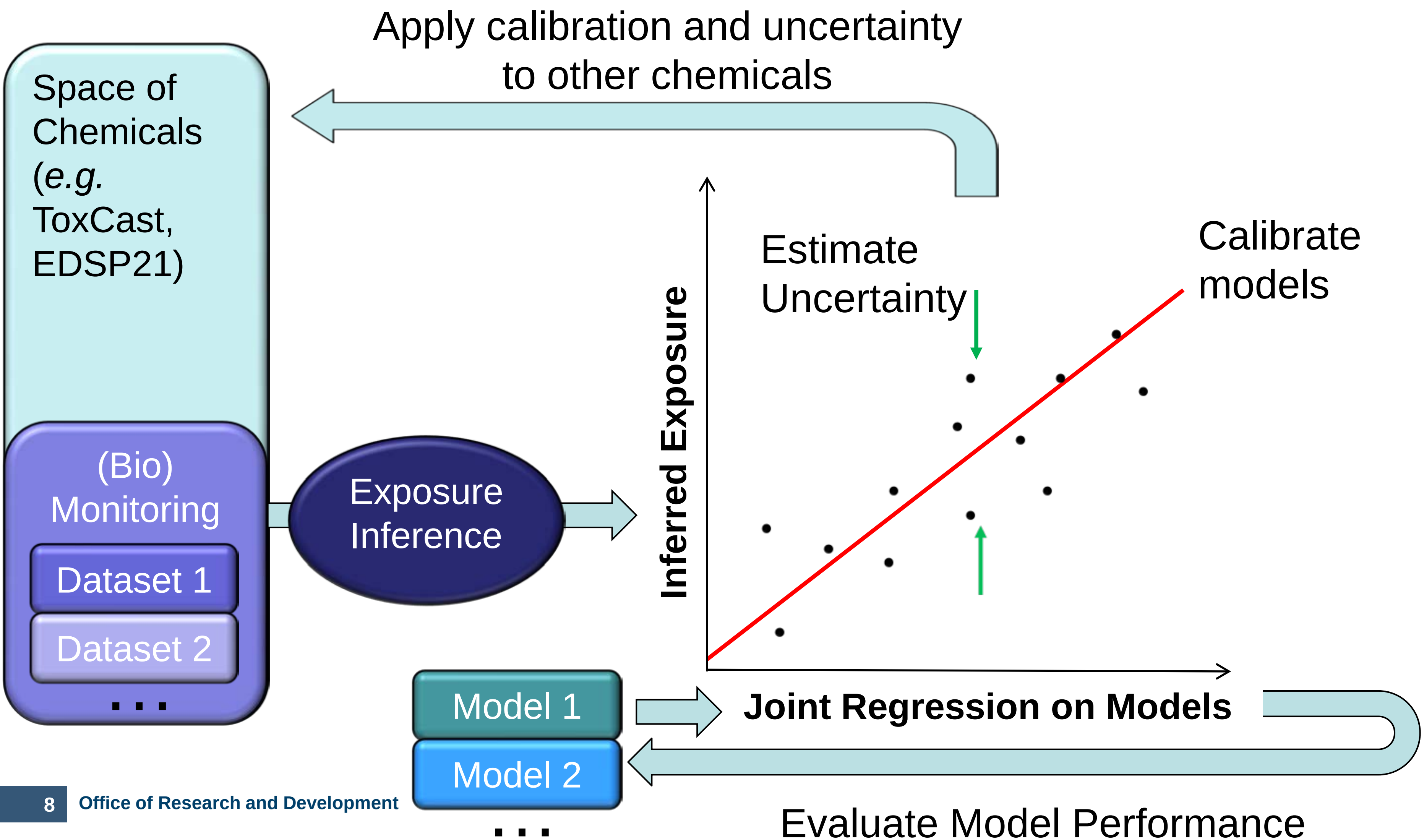
To date have found only fate and transport models to have sufficient throughput (Mitchell *et al.,* Science of the Total Environment)

Also used a simple consumer use heuristic (Frame *et al., in preparation*)

Environmental Fate and Transport

Consumer Use and Indoor Exposure

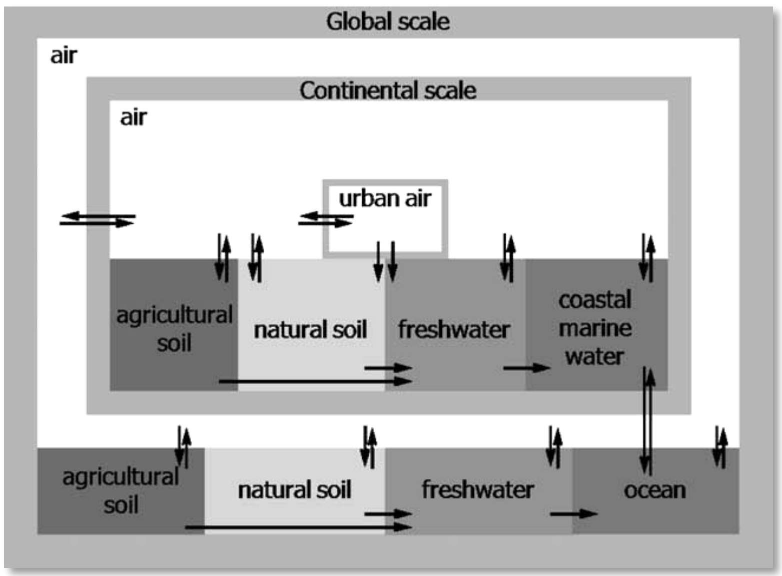
Framework for High Throughput Exposure Screening



Off the Shelf Models

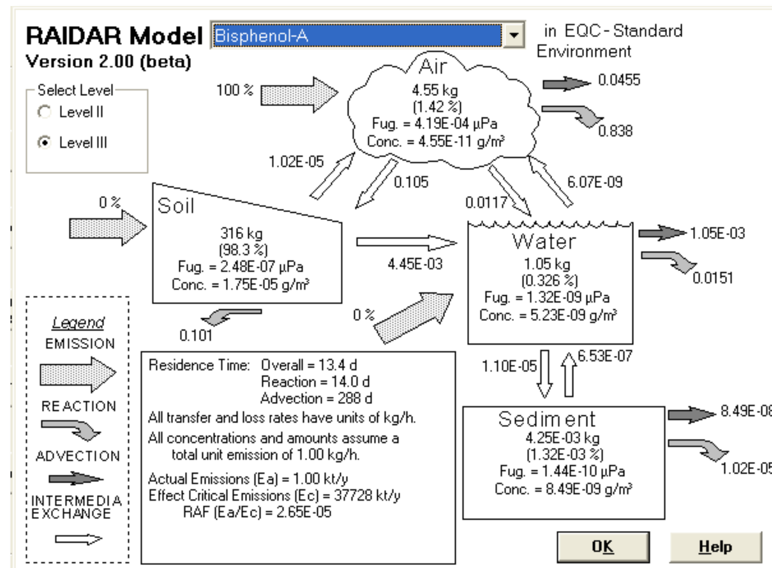
Treat different models like related high-throughput assays – consensus

USEtox



United Nations Environment Program and Society for Environmental Toxicology and Chemistry toxicity model Version 1.01
Rosenbaum *et al.* 2008

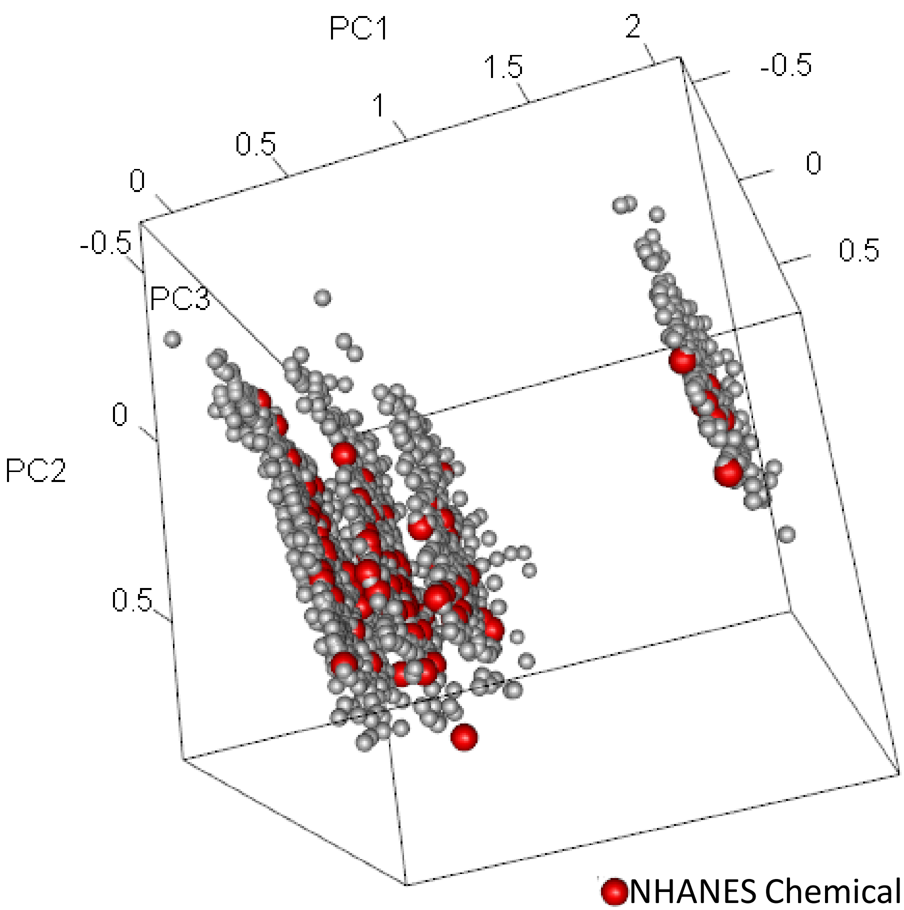
RAIDAR



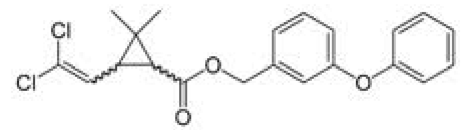
Risk Assessment IDentification And Ranking model Version 2.0
Arnot *et al*. 2006

Parameterizing the Models

Model parameters obtained from EPI Suite



Cl/C(Cl)=C/C3C(C(=O)OCc2cccc(O
c1ccccc1)c2)C3(C)C



EPI Suite contained experimental values for all parameters for ~5% of the chemicals

Many properties predicted from structure (SMILES), which failed 167 of 2127 chemicals

Dominant principal component (half life in environmental media) determined by expert elicitation

New data needed both to assess QSAR reliability and expand QSAR domain of applicability

United States Environmental Protection Agency

Data Availability for Evaluating Predictions

CDC NHANES (National Health and Nutrition Examination Survey): covers a few hundred metabolites of environmental chemicals.

Observations: parent exposures for 82 chemicals estimated by Bayesian inference based on NHANES.

- parent exposures from urinary metabolites
- focusing on U.S. total geometric mean initially

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

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

	Survey years	Geometric mean (95% conf. interval)	Selected pe (95% confiden 50th	75th
Total	03-04	**2.64** (2.38-2.94)	**2.80** (2.50-3.10)	**5.50** (5.00-6.20)
	05-06	**1.90** (1.79-2.02)	**2.00** (1.90-2.00)	**3.70** (3.50-3.90)
	07-08	**2.08** (1.92-2.26)	**2.10** (1.90-2.30)	**4.10** (3.60-4.60)
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)
	05-06	**2.86** (2.52-3.24)	**2.70** (2.30-2.90)	**5.00** (4.40-5.80)
	07-08	**2.46** (2.20-2.75)	**2.40** (1.90-3.00)	**4.50** (3.70-5.50)
12-19 years	03-04	**3.74** (3.31-4.22)	**4.30** (3.60-4.60)	**7.80** (6.50-9.00)
	05-06	**2.42** (2.18-2.68)	**2.40** (2.10-2.70)	**4.30** (3.90-5.20)
	07-08	**2.44** (2.14-2.78)	**2.30** (2.10-2.60)	**4.40** (3.70-5.50)
20 years and older	03-04	**2.41** (2.15-2.72)	**2.60** (2.30-2.80)	**5.10** (4.50-5.70)
	05-06	**1.75** (1.62-1.89)	**1.80** (1.70-2.00)	**3.40** (3.10-3.70)
	07-08	**1.99** (1.82-2.18)	**2.00** (1.80-2.30)	**3.90** (3.40-4.60)

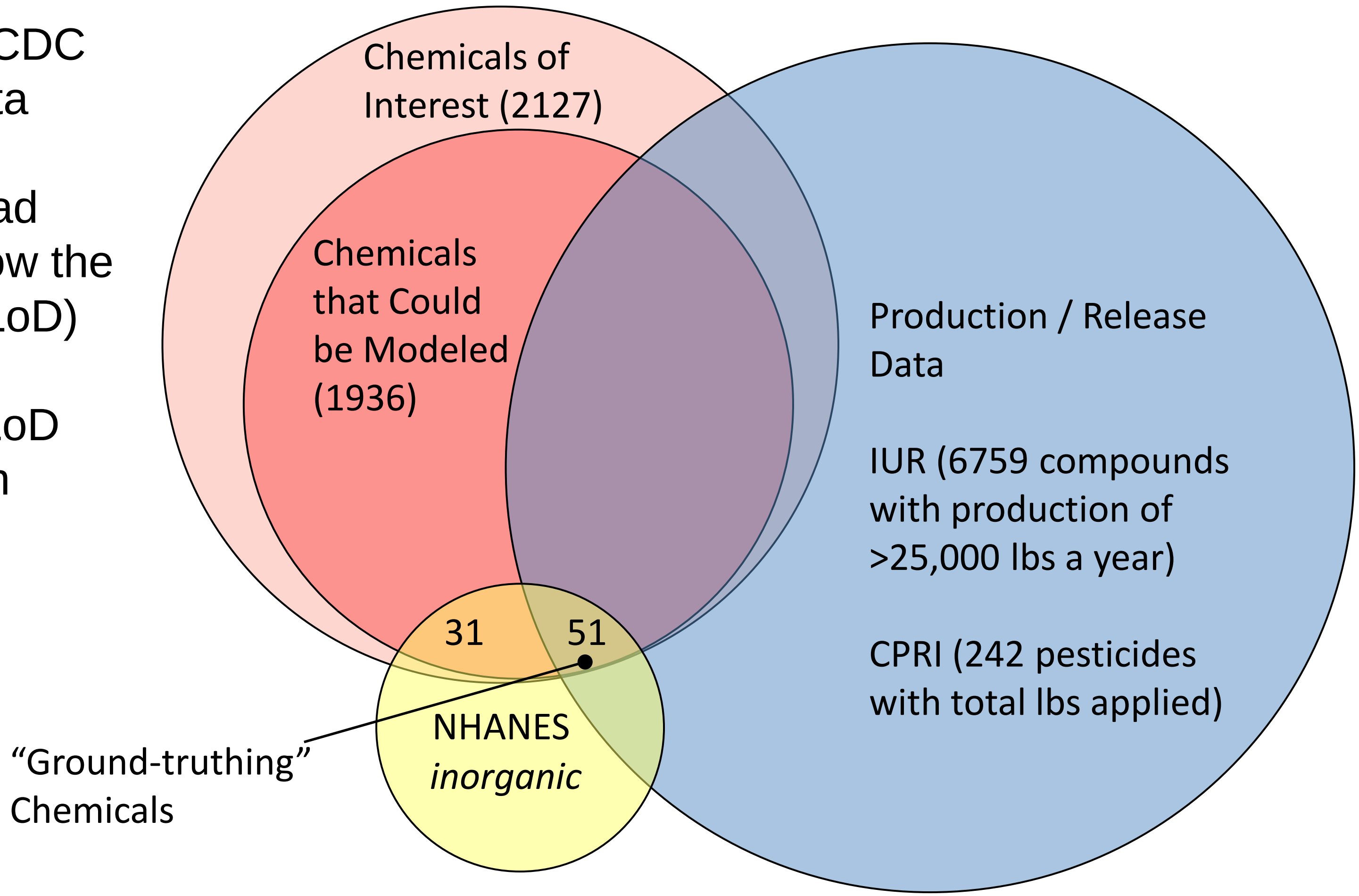
CDC, Fourth National Exposure Report （2011）

Data Availability for Model Predictions and Ground-truthing

Ground-truth with CDC NHANES urine data

Many chemicals had median conc. below the limit of detection (LoD)

Most chemicals >LoD not high production volume

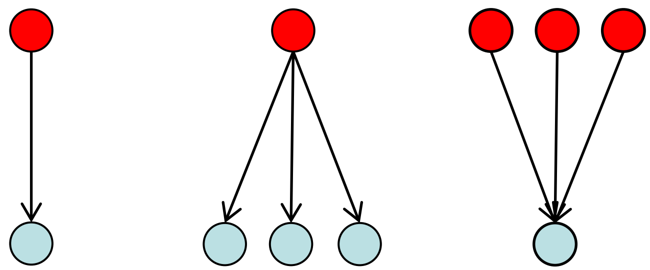


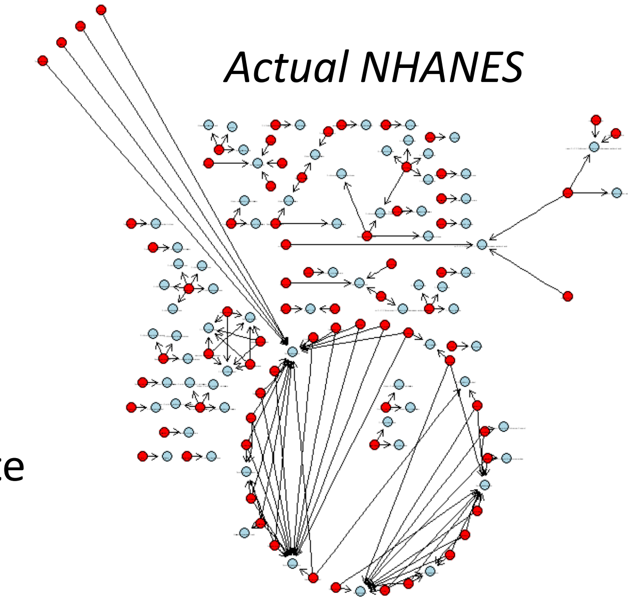
United States Environmental Protection Agency

Exposure Inference from Biomonitoring Data

Parent chemical exposure

Chemical measured In urine





A finite number of parent exposures are related to a finite number of urine products, and most of relationships are zero

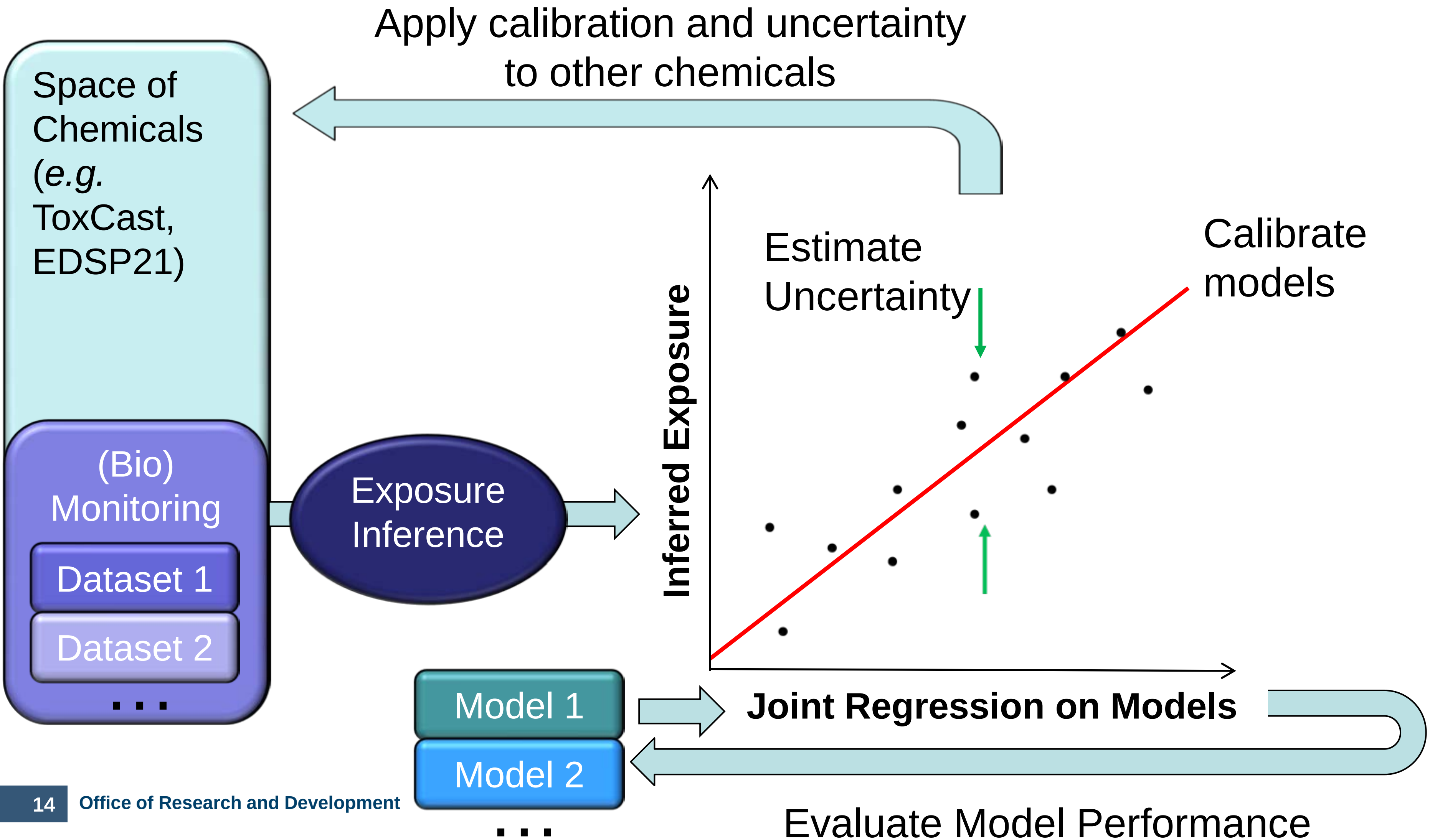
We can not determine the one "correct" combination of exposures that explains the urine concentrations for a given demographic

Use Bayesian analysis via Markov Chain Monte Carlo to create a series of different explanations that covers all likely possibilities

Separate inferences need to be done for each demographic

Work with Cory Strope, Woody Setzer

Framework for High Throughput Exposure Screening



Framework for High Throughput Exposure Screening

United States Environmental Protection Agency

Space of Chemicals (*e.g.* ToxCast, EDSP21)

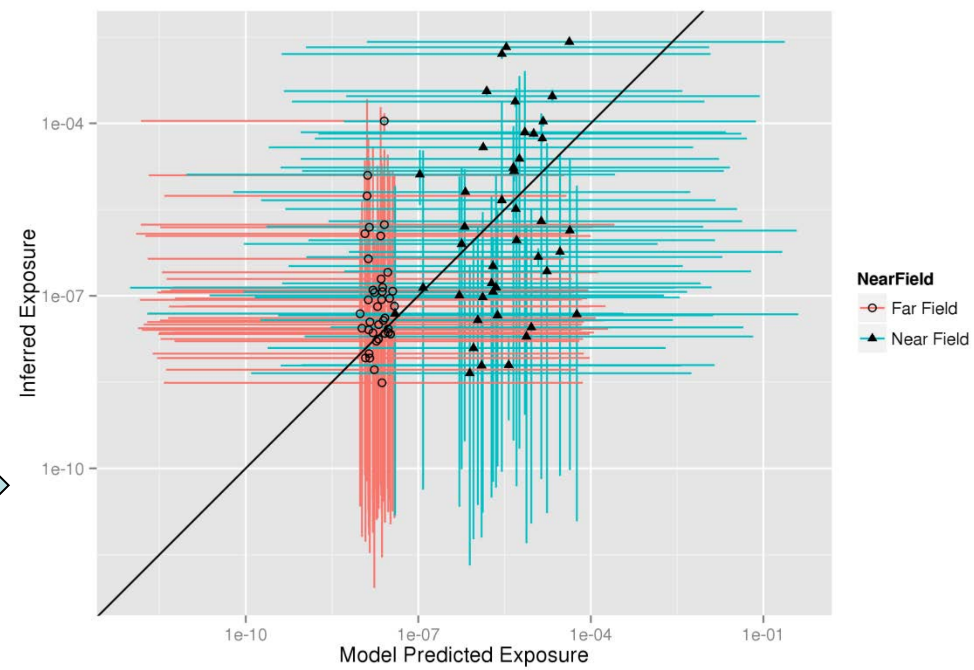
(Bio) Monitoring

Dataset 1

Dataset 2

. . .

Exposure Inference



Model 1

Model 2

. . .

Joint Regression on Models

Evaluate Model Performance

EPA United States Environmental Protection Agency

Regression on Multiple Factors

Model	Description	Mean AIC	R-squared	p-Value
0	intercept	481.98	0.00	
1	intercept + NearField	474.98	0.10	0.010
2	intercept + NearField + lV	475.67	0.12	
3	intercept + NearField + ltR	476.27	0.11	
4	intercept + NearField + luR + lV	476.41	0.14	
5	intercept + NearField + ltU	475.12	0.13	
6	intercept + NearField + luU + lV	477.08	0.13	
7	intercept + NearField + luR + luU + lV	477.24	0.15	
8	intercept + NearField + ltR + ltU	475.84	0.15	
9	intercept + NearField + NearField * luR + luU + lV	476.11	0.18	
10	intercept + NearField + luR + NearField * luU + lV	478.09	0.16	
11	intercept + NearField + luR + luU + NearField*lV	478.51	0.16	
12	intercept + NearField + NearField * luR + NearField * luU + NearField * lV	479.07	0.19	
13	intercept + NearField + NearField*luR + NearField*luU + lV	477.61	0.18	
14	intercept + NearField + luR + NearField*luU + NearField*lV	479.58	0.16	
15	intercept + NearField + NearField*luR + luU + NearField*lV	477.54	0.19	
16	intercept + NearField + NearField:luR + NearField:luU + NearField:lV	475.33	0.17	0.020
17	intercept + NearField + NearField:luR + NearField:luU	474.91	0.15	0.017
18	intercept + NearField + NearField:luR + NearField:lV	473.99	0.16	0.020
19	intercept + NearField + NearField:luU + NearField:lV	477.04	0.13	
20	intercept + NearField + NearField:luR	473.24	0.14	0.006
21	intercept + ltR	481.51	0.04	0.194
22	intercept + ltU	480.10	0.06	0.077

lV = ln(Production Volume), ltR = ln(Total RAIDAR), luR = ln(Unit RAIDAR), ltU = ln(Total USEtox), luU = ln(Unit USEtox), NearField = 0 for far-field chems, 1 for near-field, NearField * luX = separate slopes for luX, NearField : luX = slope only for NearField = 1, ltR = luR + lV, ltU = luU + lV

Analysis by Woody Setzer

Regression on Multiple Factors

Model	Description	Mean AIC	R-squared	p-Value
0	intercept	481.98	0.00	
1	intercept + NearField	474.98	0.10	0.010
17	intercept + NearField + NearField:luR + NearField:luU	474.91	0.15	0.017
21	intercept + ltR	481.51	0.04	0.194
22	intercept + ltU	480.10	0.06	0.077

lV = ln(Production Volume), ltR = ln(Total RAIDAR), luR = ln(Unit RAIDAR), ltU = ln(Total USEtox), luU = ln(Unit USEtox), NearField = 0 for far-field chems, 1 for near-field, NearField * luX = separate slopes for luX, NearField : luX = slope only for NearField = 1, ltR = luR + lV, ltU = luU + lV

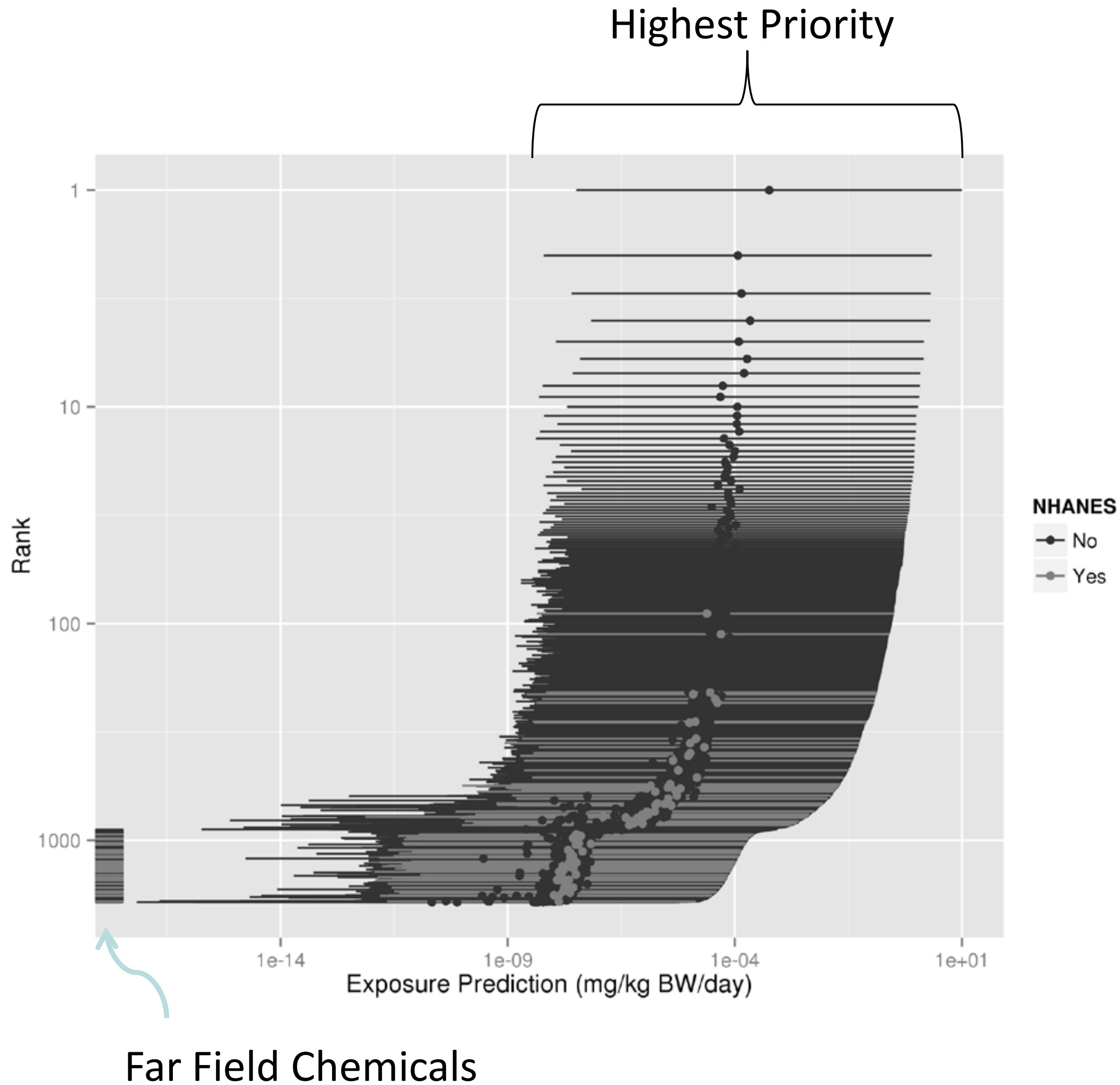
Analysis by Woody Setzer

Forecasting Exposure for 1936 Chemicals

Empirical calibration to exposures inferred from NHANES data for general population

Limited data gives broad uncertainty, but does indicate ability to forecast (R^2 = ~15%)

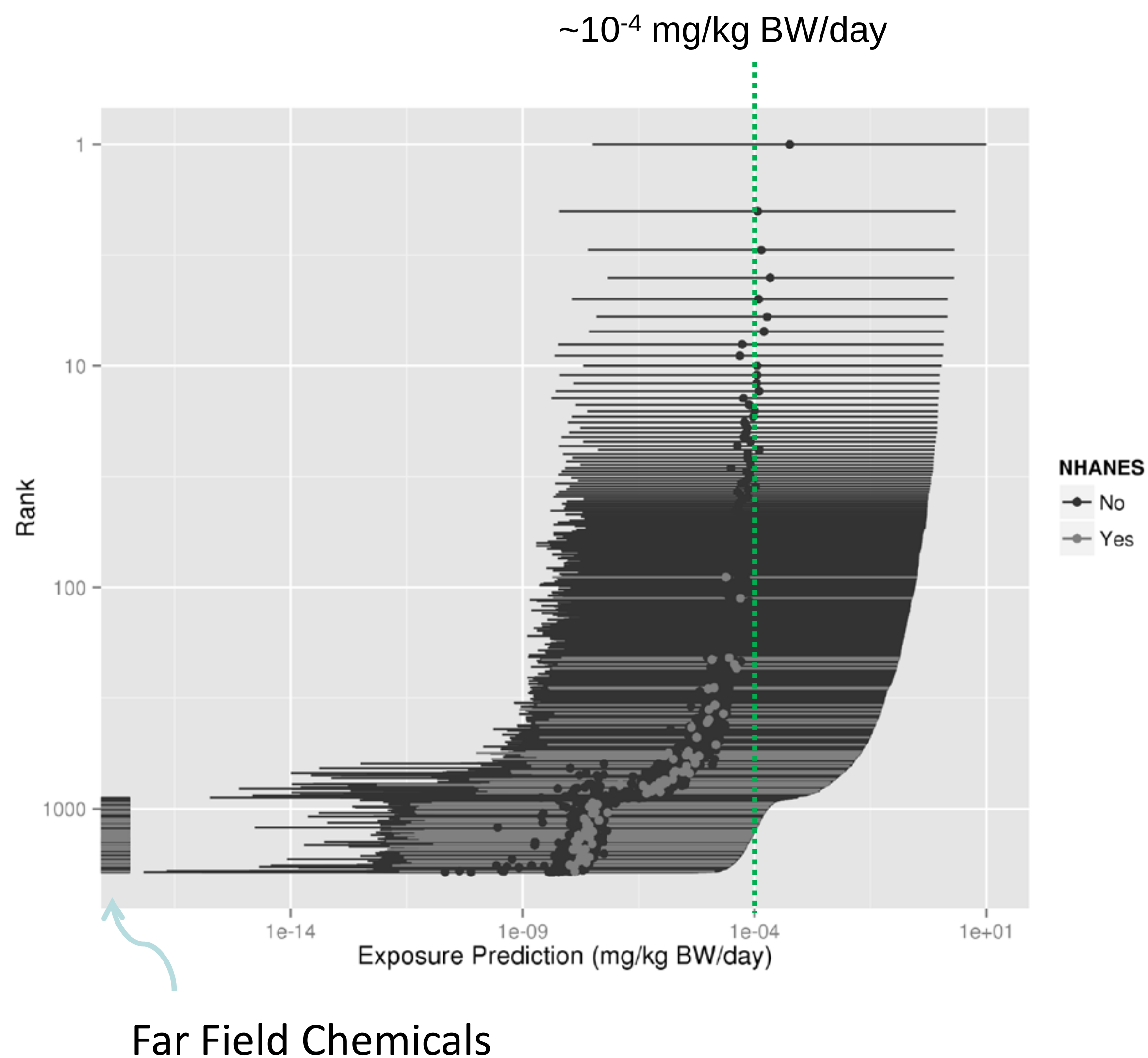
Importance of near field chemical/product use was demonstrated



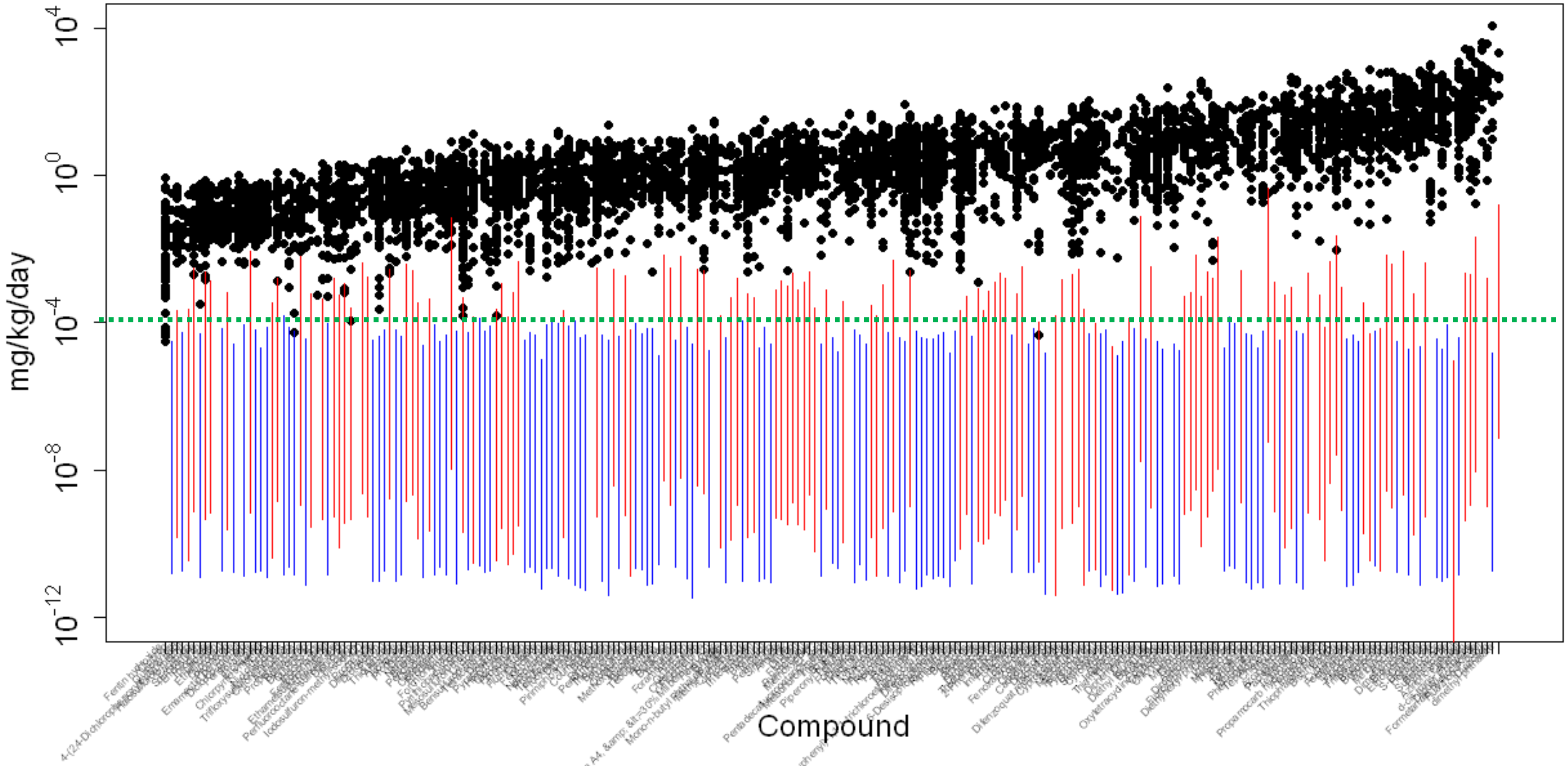
For Some Chemicals, Eight is Enough

In Wetmore *et al.* the majority doses predicted to cause ToxCast bioactivities were in excess of 10^{-4} mg/kg/day

Even with large estimated uncertainty, that the upper-limit of the 95% confidence intervals for the bottom 668 chemicals are below this level



ToxCast + ExpoCast



Oral Equivalents from Wetmore *et al.* (2012)

Paper Available from ES&T

United States
Environmental Protection
Agency

Environmental Science & Technology

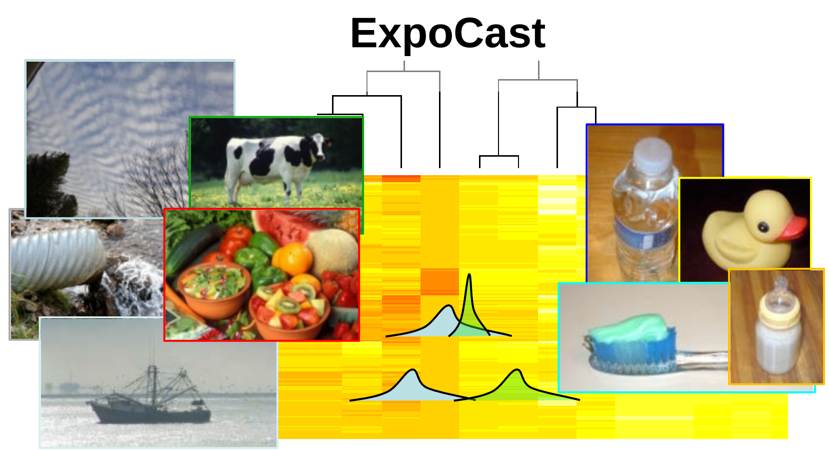
Article

High Throughput Models for Exposure-Based Chemical Prioritization in the ExpoCast Project

John F. Wambaugh, R. Woodrow Setzer, David M. Reif, Sumit Gangwal, Jade Mitchell-Blackwood, Jon A. Arnot, Olivier Joliet, Alicia Frame, James R. Rabinowitz, Thomas B. Knudsen, Richard S. Judson, Peter Egeghy, Daniel A. Vallero, and Elaine A. Cohen Hubal

Environ. Sci. Technol., **Just Accepted Manuscript** • DOI: 10.1021/es400482g • Publication Date (Web): 12 Jun 2013

Just Accepted



Statement of New Problem: Data Concerns

• If a simple near-field/far-field heuristic was most predictive so far, 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, 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

Statement of New Problem: Statistical Concerns

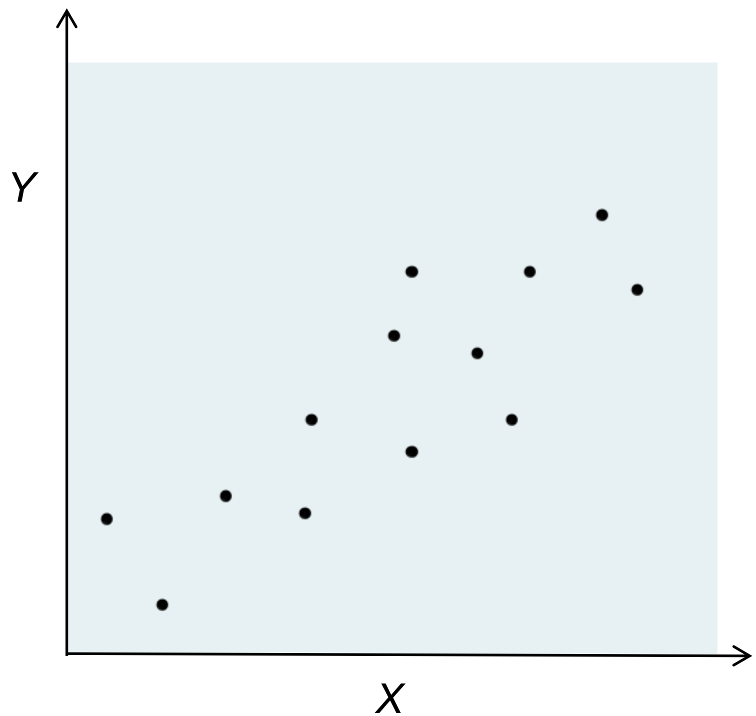
• Before we were evaluating existing models with the available (few) chemicals

• Now we are trying to build a model using essentially the same number of chemicals:

there is a danger of over-fitting

• Occam's razor (itself a heuristic) "*Plurality is not to be posited without necessity*"

• **AIC** (Akaike (1974) information criterion): the most parsimonious ("best") model has the lowest AIC score.



Noisy data and of Over-fitting

Figure from Amber Wang

Statement of New Problem: Statistical Concerns

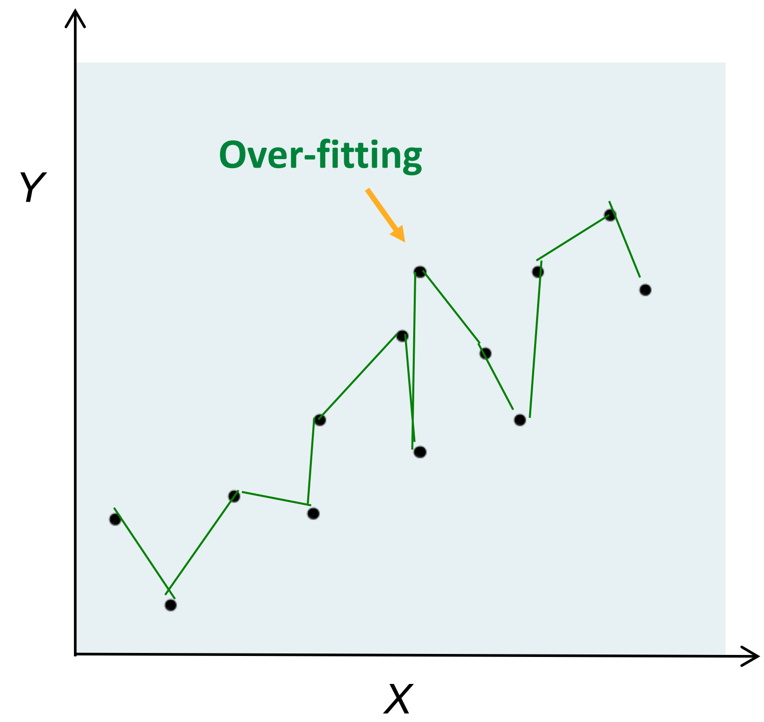
• Before we were evaluating existing models with the available (few) chemicals

• Now we are trying to build a model using essentially the same number of chemicals:

there is a danger of over-fitting

• Occam's razor (itself a heuristic) "*Plurality is not to be posited without necessity"*

• **AIC** (Akaike (1974) information criterion): the most parsimonious ("best") model has the lowest AIC score.



Noisy data and of Over-fitting

Figure from Amber Wang

Statement of New Problem: Statistical Concerns

United States Environmental Protection Agency

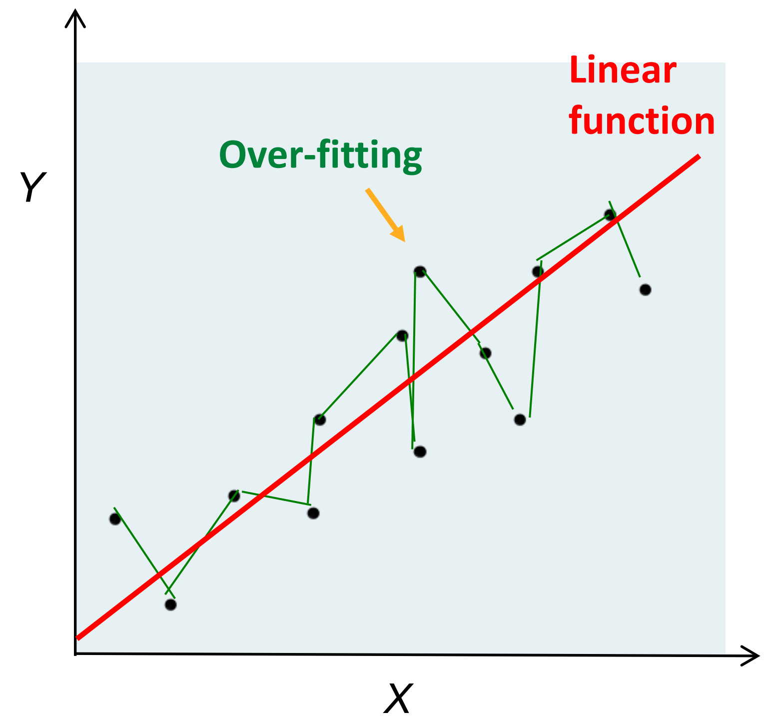
- Before we were evaluating existing models with the available (few) chemicals

- Now we are trying to build a model using essentially the same number of chemicals:

  there is a danger of over-fitting

- Occam's razor (itself a heuristic) "*Plurality is not to be posited without necessity*"

- **AIC** (Akaike (1974) information criterion): the most parsimonious ("best") model has the lowest AIC score.



Noisy data and of Over-fitting

Figure from Amber Wang

Heuristics for Chemical Use

Chemical Use Categories estimated from ACToR (chemical toxicity 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

Work by Alicia Frame, Richard Judson, slide from Amber Wang
Frame et al., *in preparation*

ExpoCast view of the NHANES (Evaluation) Chemicals

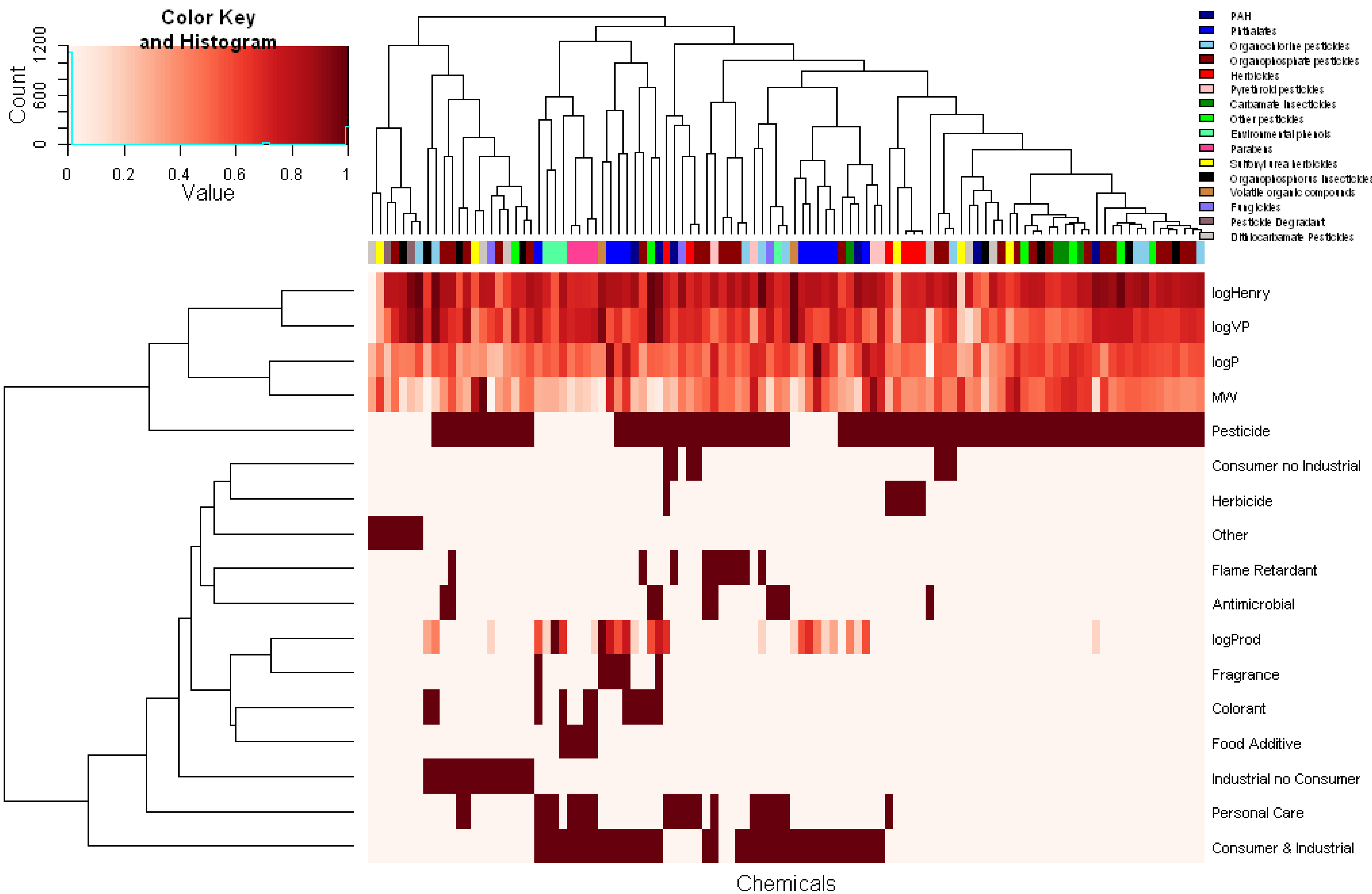


Figure from Amber Wang

EPA United States Environmental Protection Agency

Best Subset of Heuristics

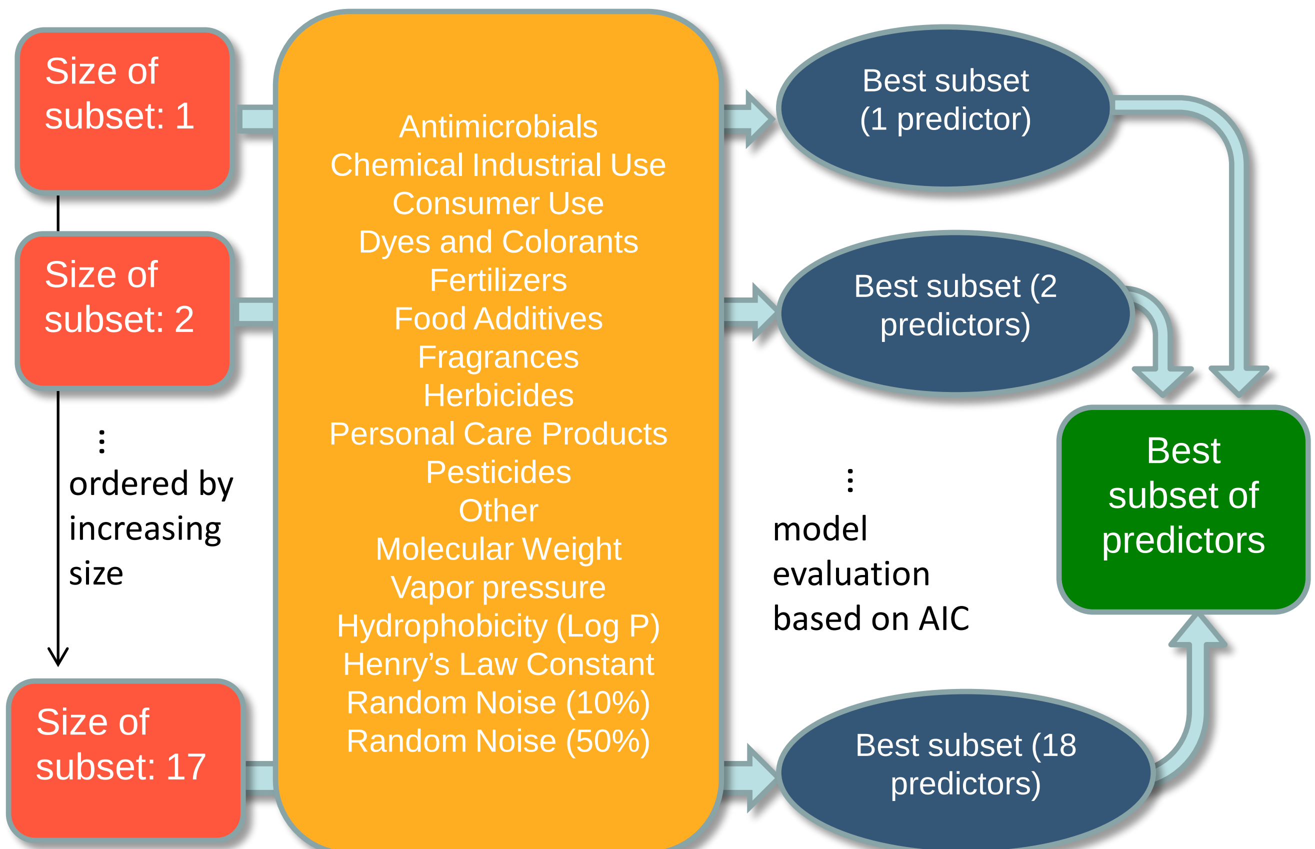
Stepwise methods search fewer combinations and rarely select the best one.

Best subsets (linear modeling algorithms): search all possible models and select the best based on some criterion.

Exhaustive search over 2^{18} models for each sample from Markov Chain

$$Y \sim \beta_0 + X_{use}\beta_{use} + x_{VP}\beta_{VP} + x_{\log P}\beta_{\log P} + x_{prod}\beta_{prod}$$

19 Candidates of Predictors



Slide and analysis by Amber Wang

Best Heuristics for General Population

United States Environmental Protection Agency

We used Bayesian methods to infer 1500 different exposure scenarios consistent with the NHANES data

We are looking for the most parsimonious explanation for the inferred exposures

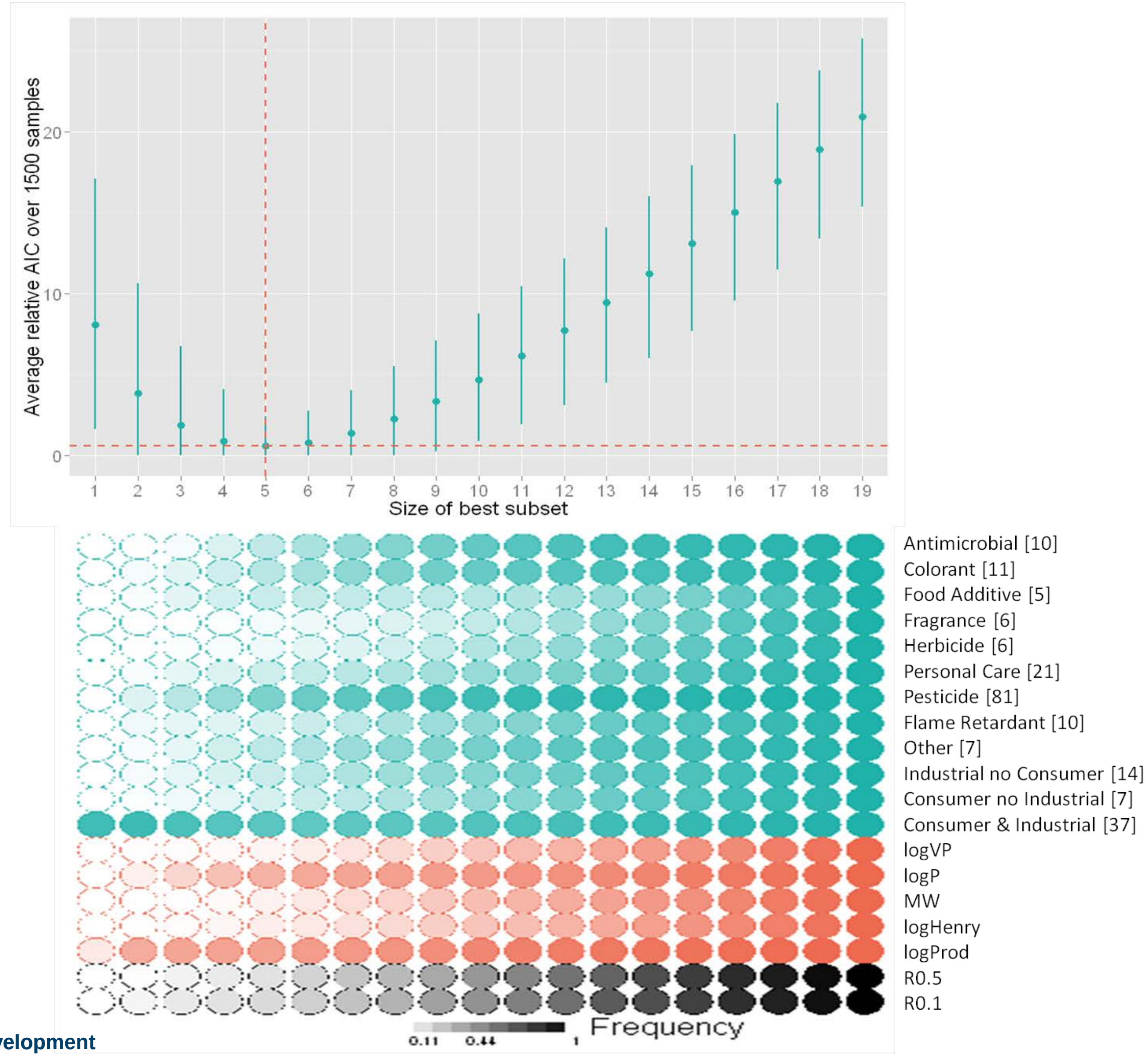


Figure from Amber Wang

NHANES Data Breaks Down by Demographics

United States Environmental Protection Agency

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

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

	Survey years	Geometric mean (95% conf. interval)	Selected pe (95% confiden 50th	75th
Total	03-04	**2.64** (2.38-2.94)	**2.80** (2.50-3.10)	**5.50** (5.00-6.20)
	05-06	**1.90** (1.79-2.02)	**2.00** (1.90-2.00)	**3.70** (3.50-3.90)
	07-08	**2.08** (1.92-2.26)	**2.10** (1.90-2.30)	**4.10** (3.60-4.60)
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)
	05-06	**2.86** (2.52-3.24)	**2.70** (2.30-2.90)	**5.00** (4.40-5.80)
	07-08	**2.46** (2.20-2.75)	**2.40** (1.90-3.00)	**4.50** (3.70-5.50)
12-19 years	03-04	**3.74** (3.31-4.22)	**4.30** (3.60-4.60)	**7.80** (6.50-9.00)
	05-06	**2.42** (2.18-2.68)	**2.40** (2.10-2.70)	**4.30** (3.90-5.20)
	07-08	**2.44** (2.14-2.78)	**2.30** (2.10-2.60)	**4.40** (3.70-5.50)
20 years and older	03-04	**2.41** (2.15-2.72)	**2.60** (2.30-2.80)	**5.10** (4.50-5.70)
	05-06	**1.75** (1.62-1.89)	**1.80** (1.70-2.00)	**3.40** (3.10-3.70)
	07-08	**1.99** (1.82-2.18)	**2.00** (1.80-2.30)	**3.90** (3.40-4.60)

CDC, Fourth National Exposure Report (2011)

- Will different demographics have different heuristics?

United States Environmental Protection Agency

NHANES Data Breaks Down by 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. and Nutrition Examination Survey.

	Survey years	Geometric mean (95% conf. interval)	Selected pe (95% confiden 50th	75th
Total	03-04	**2.64** (2.38-2.94)	**2.80** (2.50-3.10)	**5.50** (5.00-6.20)
	05-06	**1.90** (1.79-2.02)	**2.00** (1.90-2.00)	**3.70** (3.50-3.90)
	07-08	**2.08** (1.92-2.26)	**2.10** (1.90-2.30)	**4.10** (3.60-4.60)
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)
	05-06	**2.86** (2.52-3.24)	**2.70** (2.30-2.90)	**5.00** (4.40-5.80)
	07-08	**2.46** (2.20-2.75)	**2.40** (1.90-3.00)	**4.50** (3.70-5.50)
12-19 years	03-04	**3.74** (3.31-4.22)	**4.30** (3.60-4.60)	**7.80** (6.50-9.00)
	05-06	**2.42** (2.18-2.68)	**2.40** (2.10-2.70)	**4.30** (3.90-5.20)
	07-08	**2.44** (2.14-2.78)	**2.30** (2.10-2.60)	**4.40** (3.70-5.50)
20 years and older	03-04	**2.41** (2.15-2.72)	**2.60** (2.30-2.80)	**5.10** (4.50-5.70)
	05-06	**1.75** (1.62-1.89)	**1.80** (1.70-2.00)	**3.40** (3.10-3.70)
	07-08	**1.99** (1.82-2.18)	**2.00** (1.80-2.30)	**3.90** (3.40-4.60)

CDC, Fourth National Exposure Report (2011)

- Will different demographics have different heuristics?

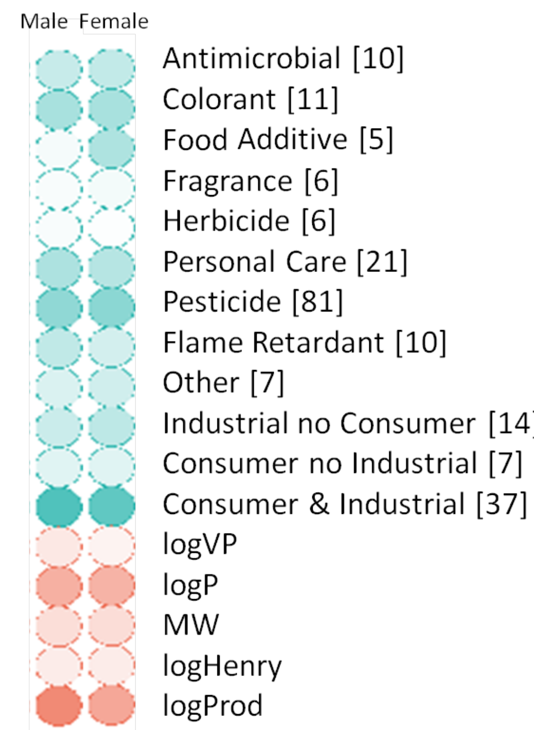


Figure from Amber Wang

Tox21 Exposure Predictions for the General U.S. Population

New empirical calibration to exposures inferred from NHANES data for general population

Reduced uncertainty from previous model

95% confidence intervals still contain most chemicals

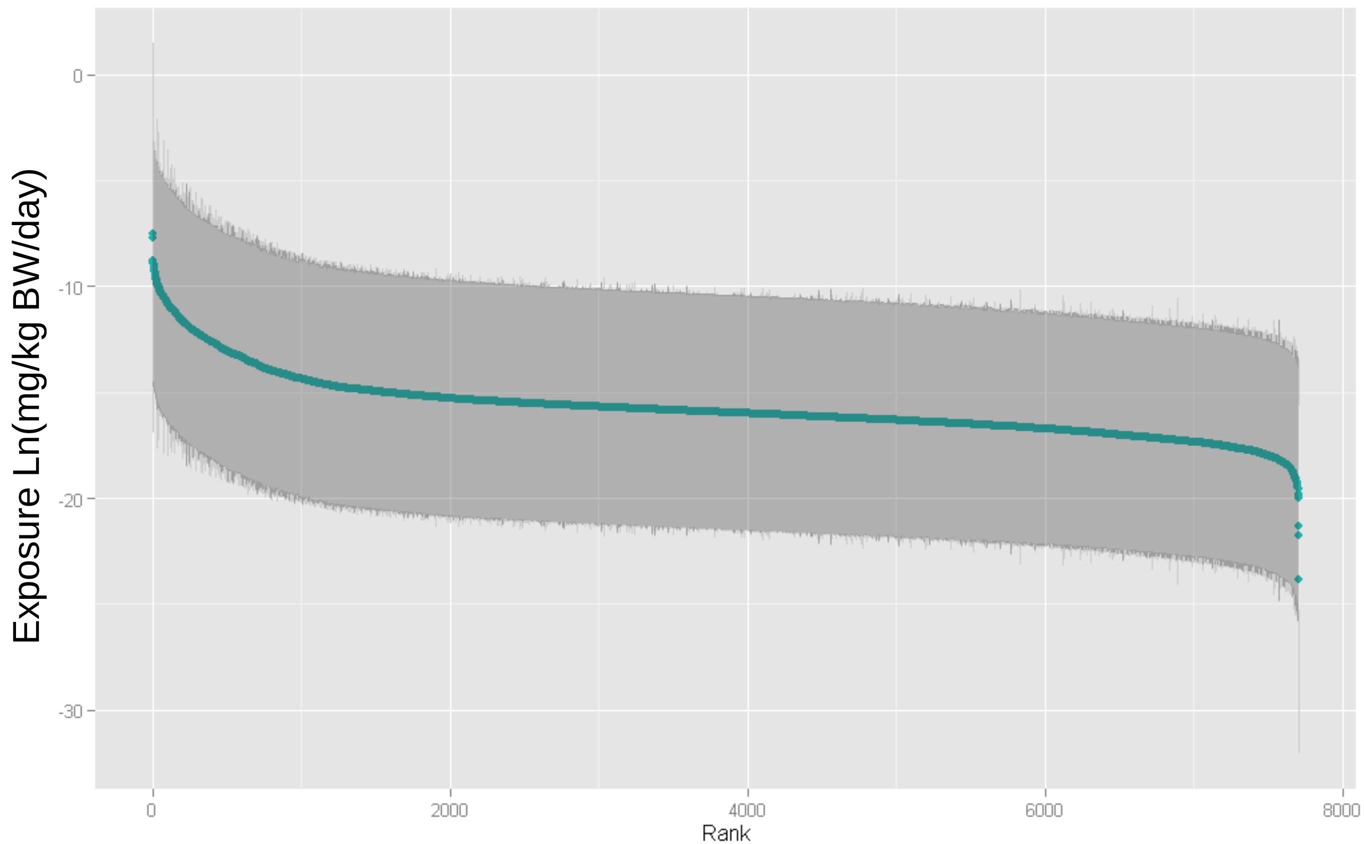


Figure from Amber Wang

Exposure Priorities

Obtaining new chemical data

- Measuring physico-chemical parameters
  - Characterizing QSAR appropriateness
  - Expanding QSAR domain of applicability
- Determining occurrence in articles, packaging, and products

New monitoring data

- Validation of predictions
- Characterization of chemical exposure
  - Specific demographics
  - Pooled (average) samples

New indoor/consumer use models

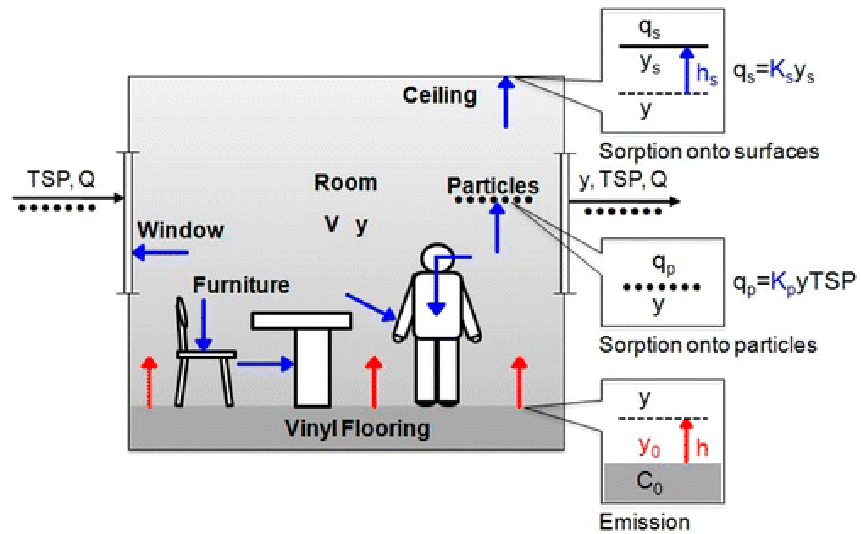


Image from Little *et al.* (2012),

see also Nazaroff *et al.* (2012), Shin *et al.* (2012), Wenger and Jolliet (2012)

Conclusions

- High throughput computational model predictions of exposure is possible
  - These prioritizations have been compared with CDC NHANES data, yielding empirical calibration and estimate of uncertainty
- Indoor/consumer use is a primary determinant of NHANES exposure
  - Developing HT models for exposure from consumer use and indoor environment (*post-doc position available*)
- Can develop demographic-specific prioritizations
- Enhanced use data (ACToR annotation and MSDS curation) available upon publication via ACToR – http://www.epa.gov/actor/

ExpoCast Opportunities

Request for Proposals (Contract):

Exposure Screening Tools for Accelerated Chemical Prioritization, SOL-NC-13-00017
http://www.epa.gov/oamrtpnc/1300017/index.htm

Post-Doctoral Research Positions:

High Throughput Pharmacokinetic Modeling of Environmental Chemicals
http://orise.orau.gov/epa/description.aspx?JobId=12640

High Throughput Modeling of Indoor Exposure to Chemicals
http://orise.orau.gov/epa/description.aspx?JobId=12641

EPA Science to Achieve Results (STAR) Grants:

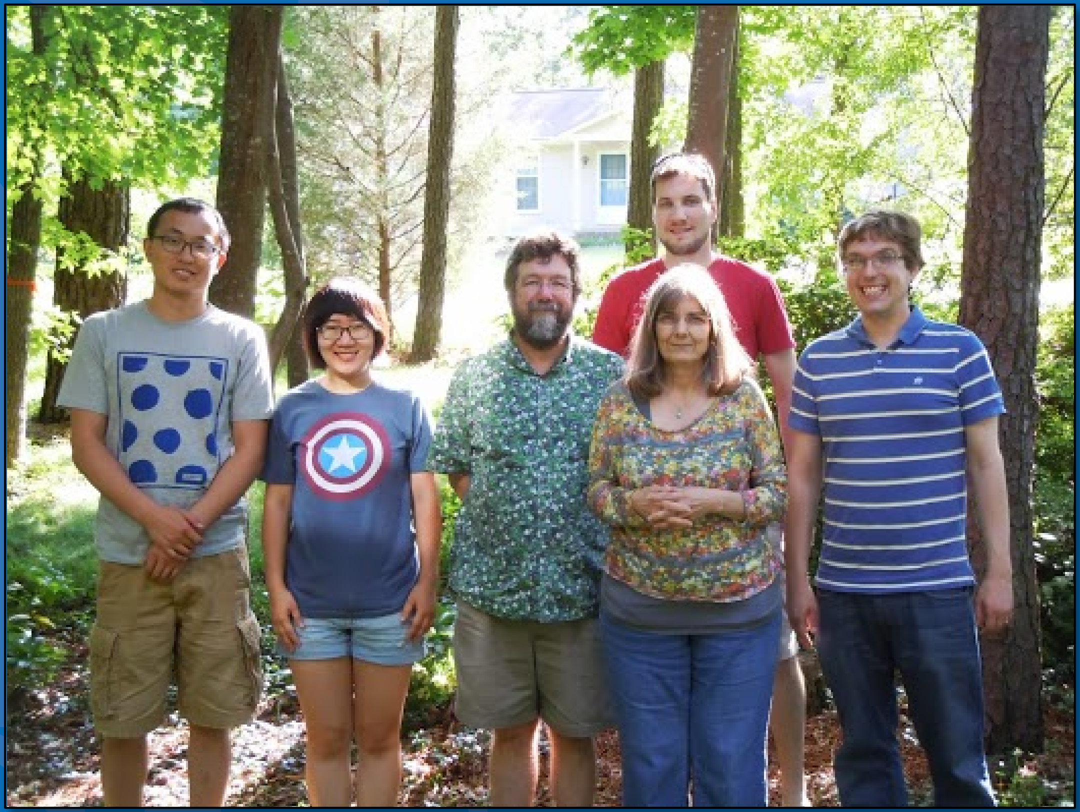
New Methods in 21st Century Exposure Science
http://epa.gov/ncer/rfa/2013/2013_star_exposure_science.html

Susceptibility and Variability in Human Response to Chemical Exposures
http://epa.gov/ncer/rfa/2013/2013_star_chemical_exposure.html

EPA Office of Research and Development

ExpoCast Team

Kathie Dionisio*
Peter Egeghy
Richard Judson
Thomas Knudsen
James Rabinowitz
Woody Setzer
Cory Strope*
Dan Vallero
Amber Wang*



***Graduate Student / Post-Doc**

External Collaborators

Jon Arnot (ARC)
Deborah Bennett (University of California, Irvine)
Alicia Frame (Dow Chemical)
Olivier Jolliet (University of Michigan)
Jade Mitchell (Michigan State)
Barbara Wetmore (Hamner)

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