United States Environmental Protection Agency

Exposure-based Chemical Priority Setting in the 21st Century

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

Duke Risk Assessment Course
March 18, 2019

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

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

EPA Office of Research and Development

United States Environmental Protection Agency

- The Office of Research and Development (ORD) is the scientific research arm of EPA
  - 626 peer-reviewed journal articles in 2017 and 562 so far for 2018
- Research is conducted by ORD's three national laboratories, four national centers, and two offices organized to address:
  - Hazard, exposure, risk assessment, and risk management
- 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
- A tapestry of laws covers the chemicals people are exposed to in the United States (Breyer, 2009)
- Different testing requirements exist for food additives, pharmaceuticals, and pesticide active ingredients (NRC, 2007)

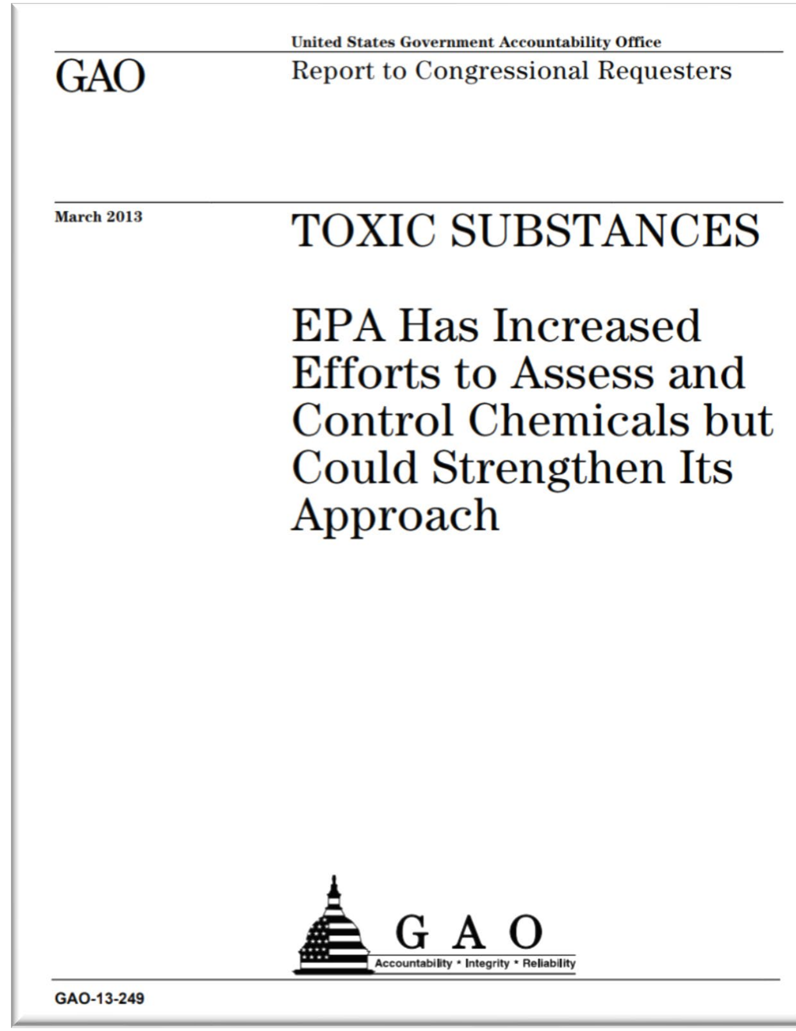
November 29, 2014

United States Environmental Protection Agency

Chemical Regulation in the United States

- Most other chemicals, ranging from industrial waste to dyes to packing materials, are covered by the Toxic Substances Control Act (TSCA)
  - Thousands of chemicals on the market were either "grandfathered" in or were allowed without experimental assessment of hazard, toxicokinetics, or exposure (Judson et al. (2009), Egeghy et al. (2012), Wetmore et al. (2015))

"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


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

G A O

Accountability • Integrity • Reliability

GAO-13-249

March, 2013

Chemical Regulation in the United States

- 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
  - "Strategic Plan to Promote the Development and Implementation of Alternative Test Methods Within the TSCA Program" (June 22, 2018)

130 STAT. 448 PUBLIC LAW 114–182—JUNE 22, 2016

Public Law 114–182
114th Congress

An Act

June 22, 2016
[H.R. 2576]

To modernize the Toxic Substances Control Act, and for other purposes.

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,

Frank R. Lautenberg Chemical Safety for the 21st Century Act. 15 USC 2601 note.

SECTION 1. SHORT TITLE; TABLE OF CONTENTS.

(a) SHORT TITLE.—This Act may be cited as the "Frank R. Lautenberg Chemical Safety for the 21st Century Act".
(b) TABLE OF CONTENTS.—The table of contents of this Act is as follows:

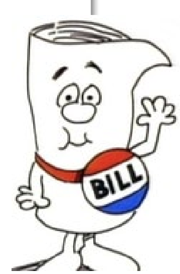
TITLE I—CHEMICAL SAFETY

SEC. 2. FINDINGS, POLICY, AND INTENT.

Section 2(c) of the Toxic Substances Control Act (15 U.S.C. 2601(c)) is amended by striking "proposes to take" and inserting "proposes as provided".

SEC. 3. DEFINITIONS.

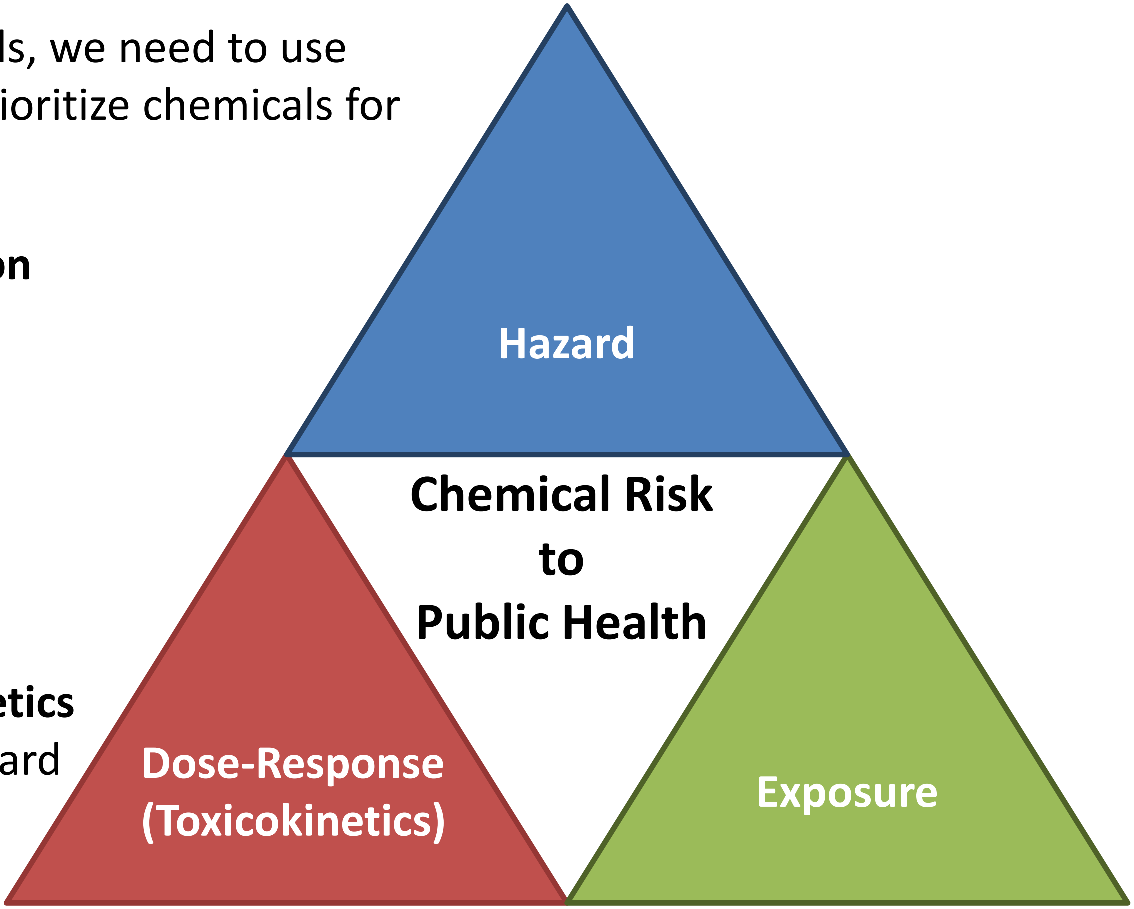
Section 3 of the Toxic Substances Control Act (15 U.S.C. 2602) is amended—



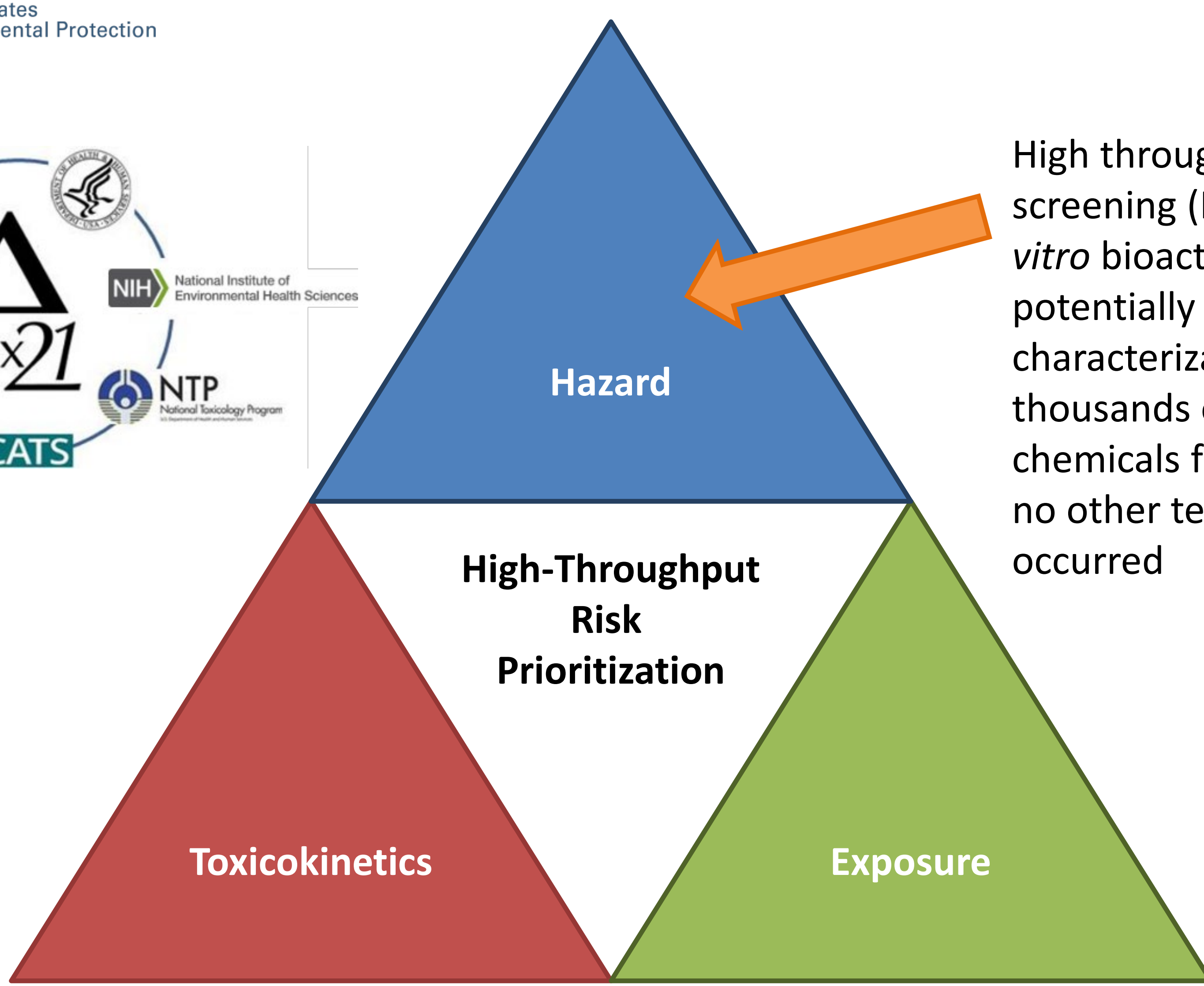
June 22, 2016

Chemical Risk = Hazard x Exposure

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



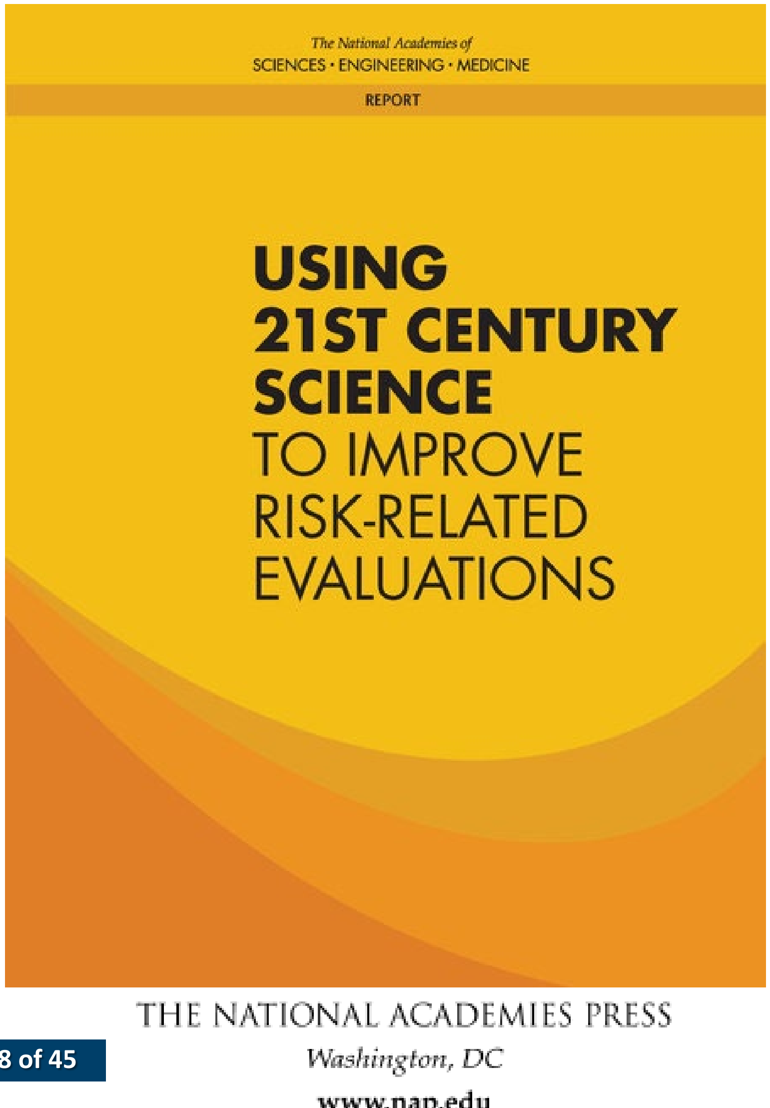
High-Throughput Risk Prioritization



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

United States Environmental Protection Agency

Risk Assessment in the 21st Century



January 5, 2017

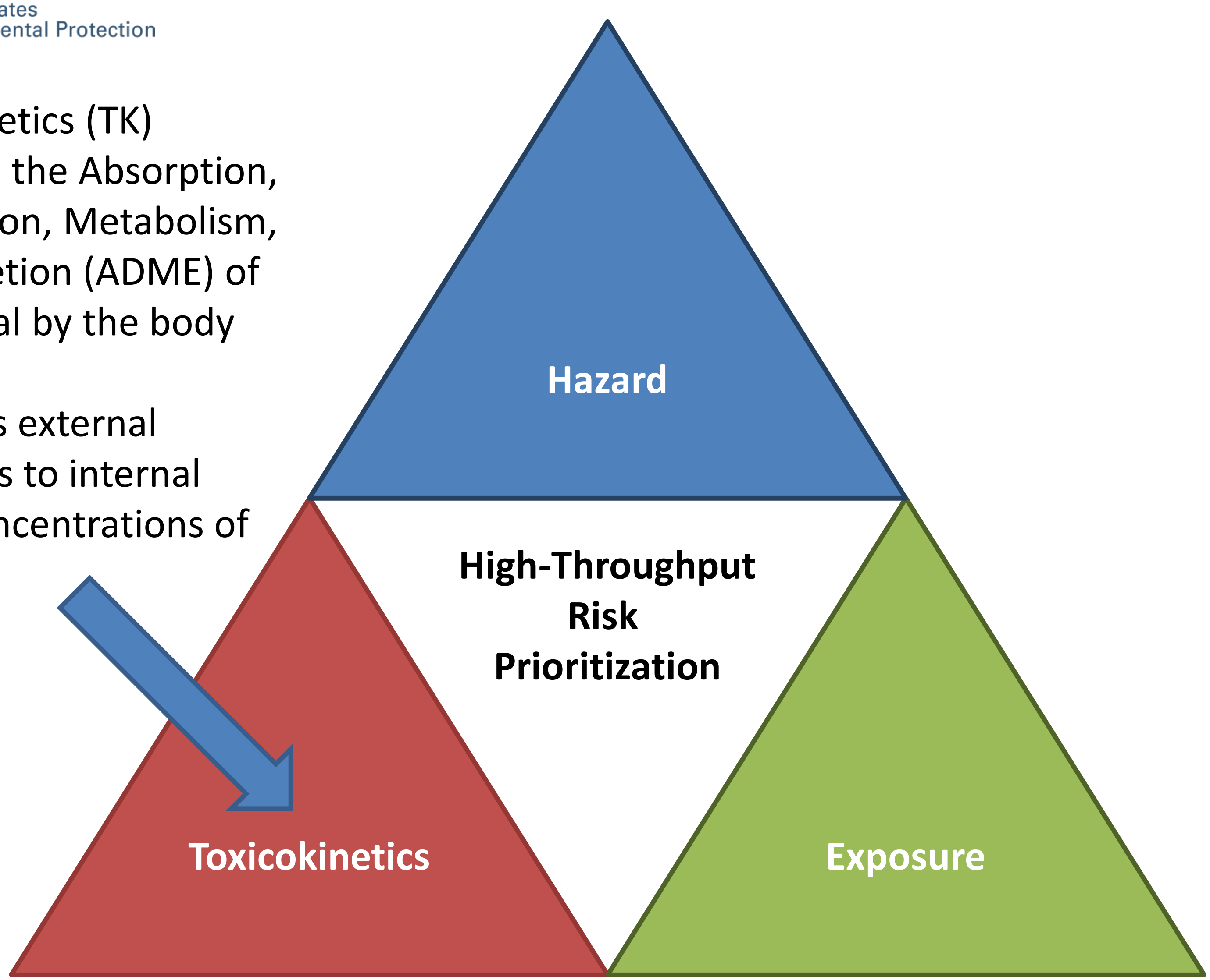
"Translation of high-throughput data into risk-based rankings is an important application of exposure data for chemical priority-setting. Recent advances in high-throughput toxicity assessment, notably the ToxCast and Tox21 programs (see Chapter 1), and in high-throughput computational exposure assessment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

"…The committee sees the potential for the application of **computational exposure science** to be highly valuable and credible for comparison and **priority-setting among chemicals in a risk-based context**."

High Throughput Toxicokinetics (HTTK)

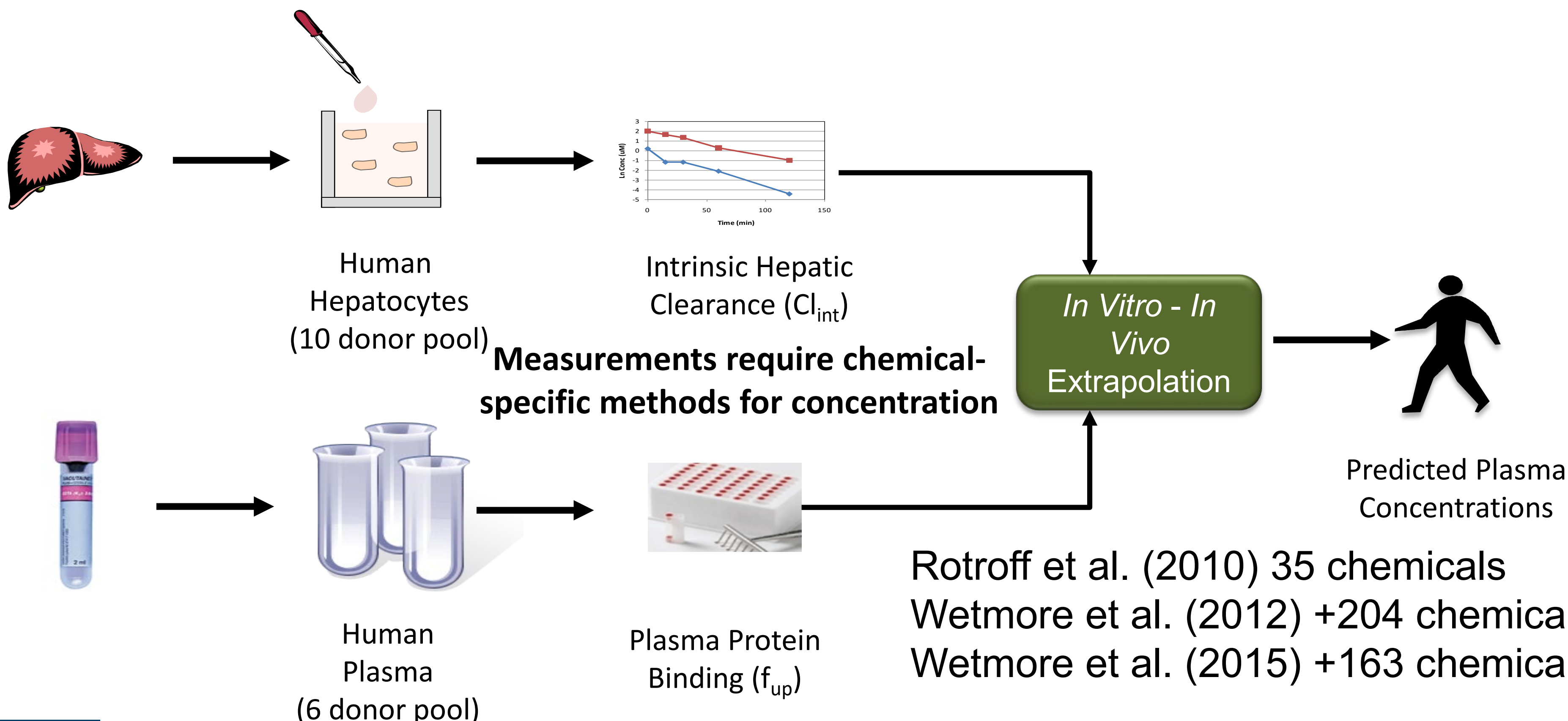
Toxicokinetics (TK) describes the Absorption, Distribution, Metabolism, and Excretion (ADME) of a chemical by the body

TK relates external exposures to internal tissue concentrations of chemical



High-Throughput Toxicokinetics (HTTK)

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



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

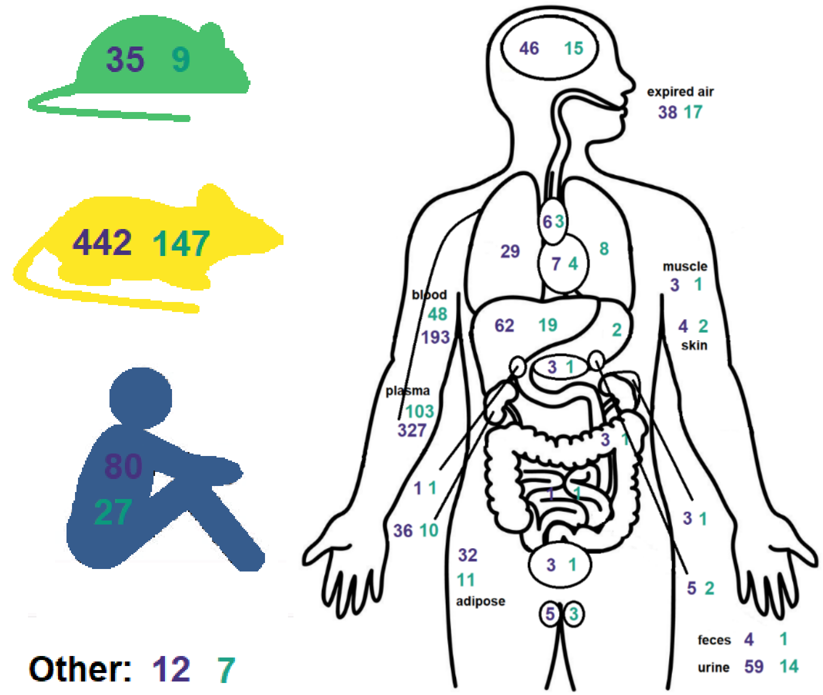
Figure from Barbara Wetmore

Building Confidence in HTTK: The Need for Data

- EPA is developing a public database of concentration vs. time data for building, calibrating, and evaluating TK models

- Curation and development ongoing, but to date includes:
  - 198 analytes (EPA, National Toxicology Program, literature)
  - Routes: Intravenous, dermal, oral, sub-cutaneous, and inhalation exposure

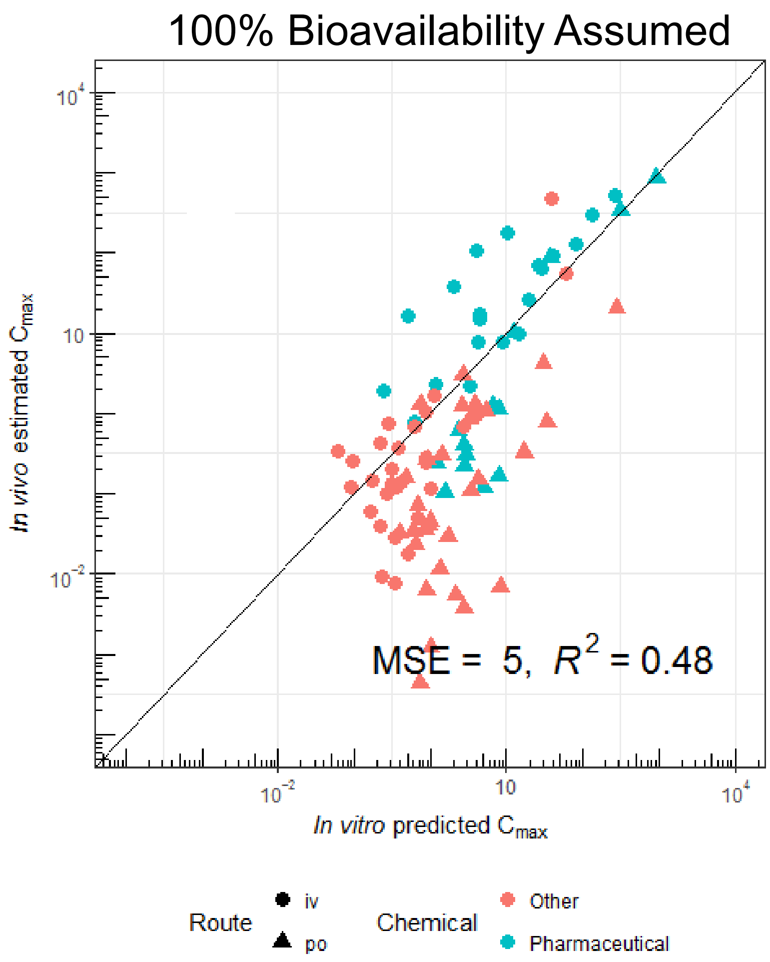
- Database will be made available through web interface and through the “httk” R package



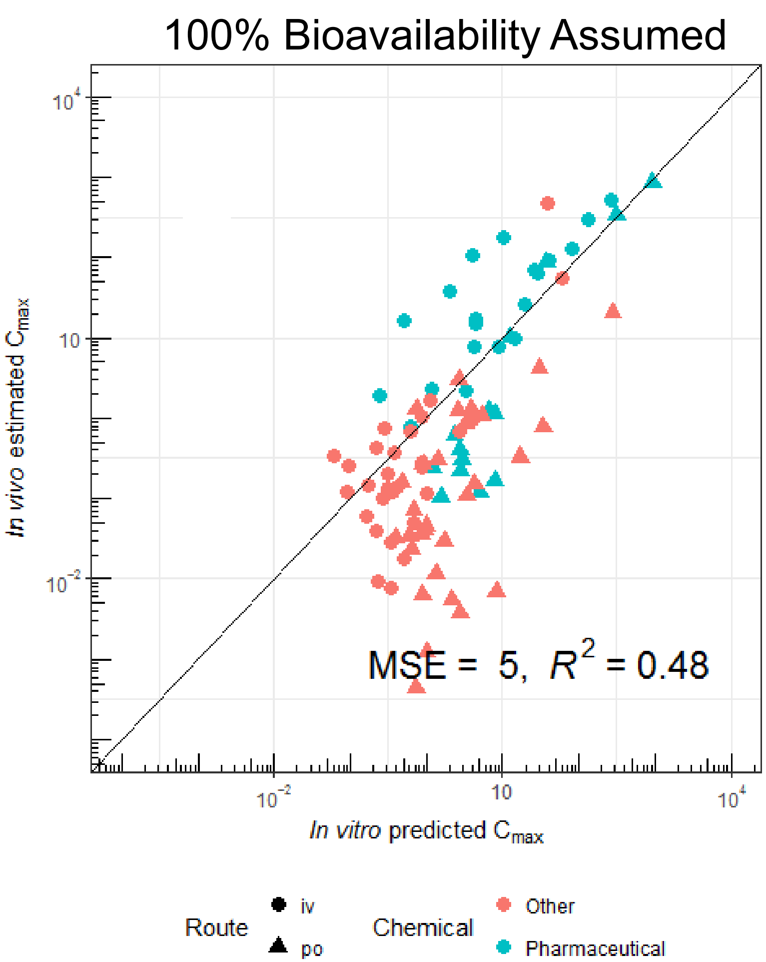
- Standardized, open source curve fitting software invivoPKfit used to calibrate models to all data: https://github.com/USEPA/CompTox-ExpoCast-invivoPKfit

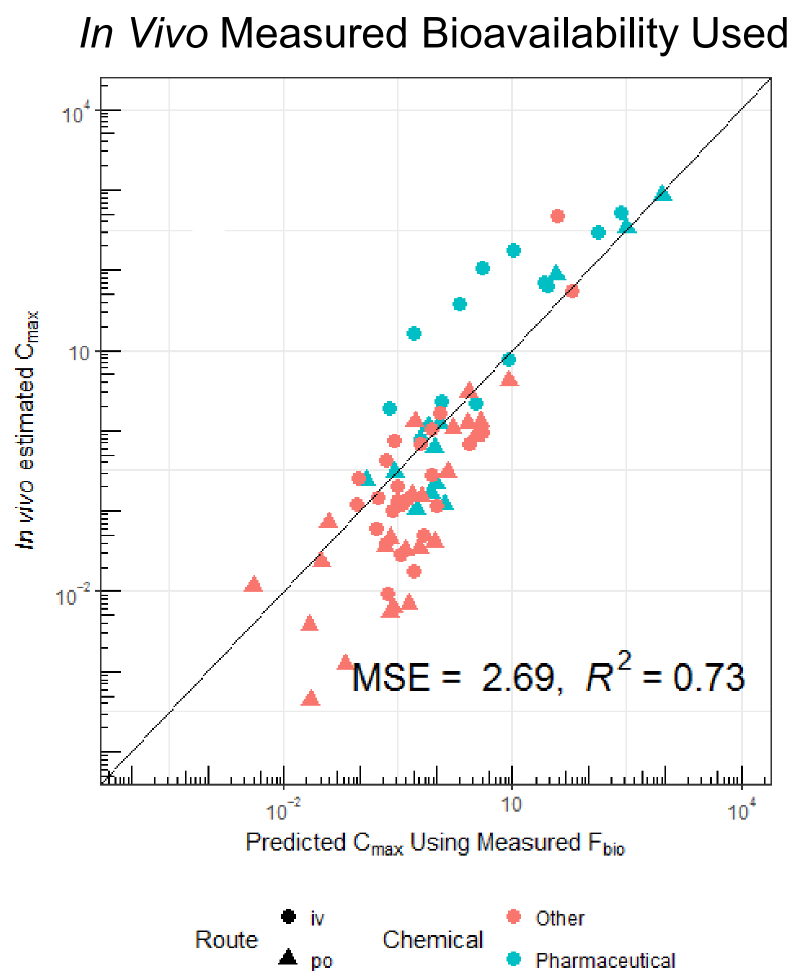
Sayre et al. (in preparation)

Evaluating HTTK



Evaluating HTTK





Here, we find that need to predict oral absorption

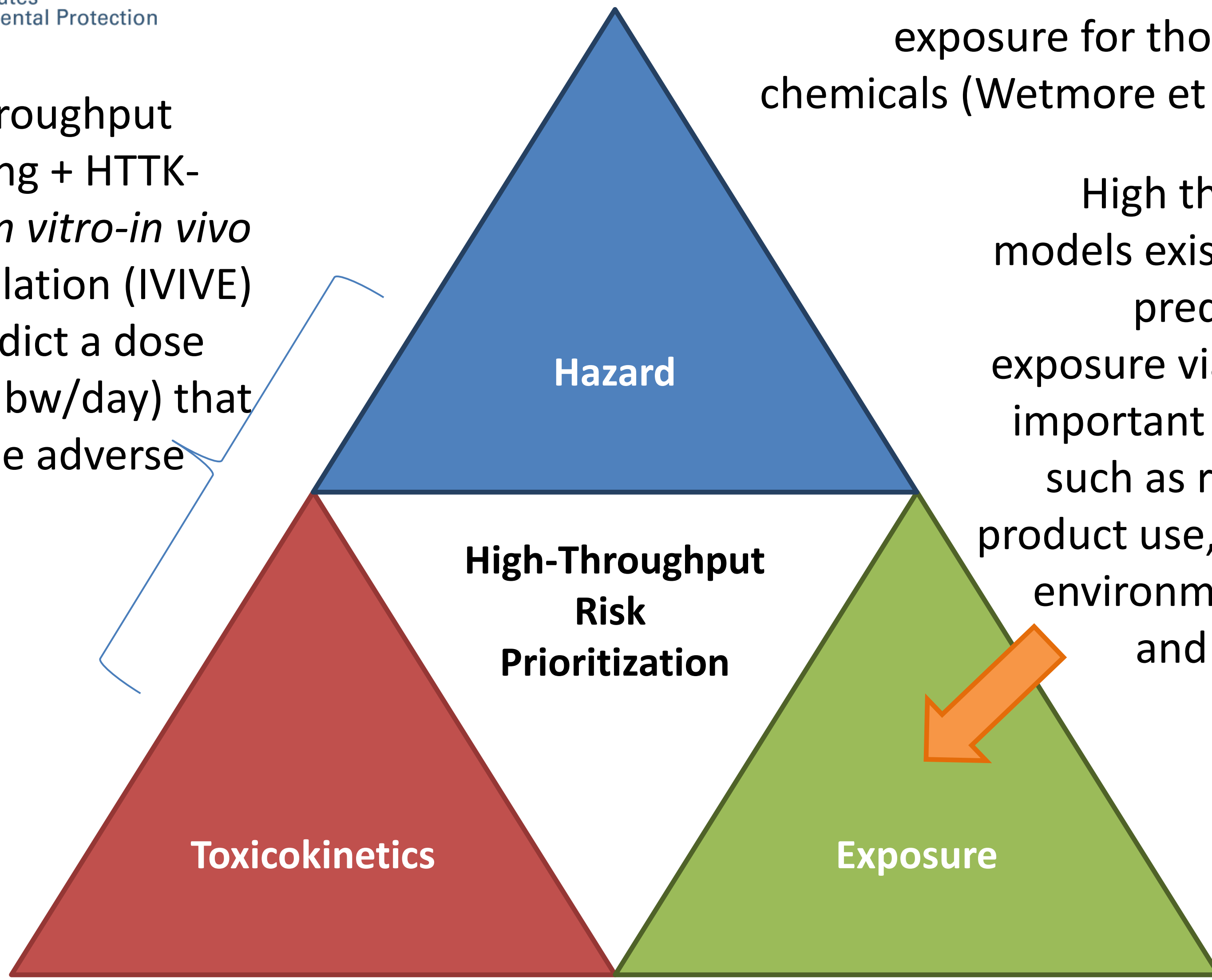
Honda et al. (in preparation)

New Exposure Data and Models

United States
Environmental Protection
Agency

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

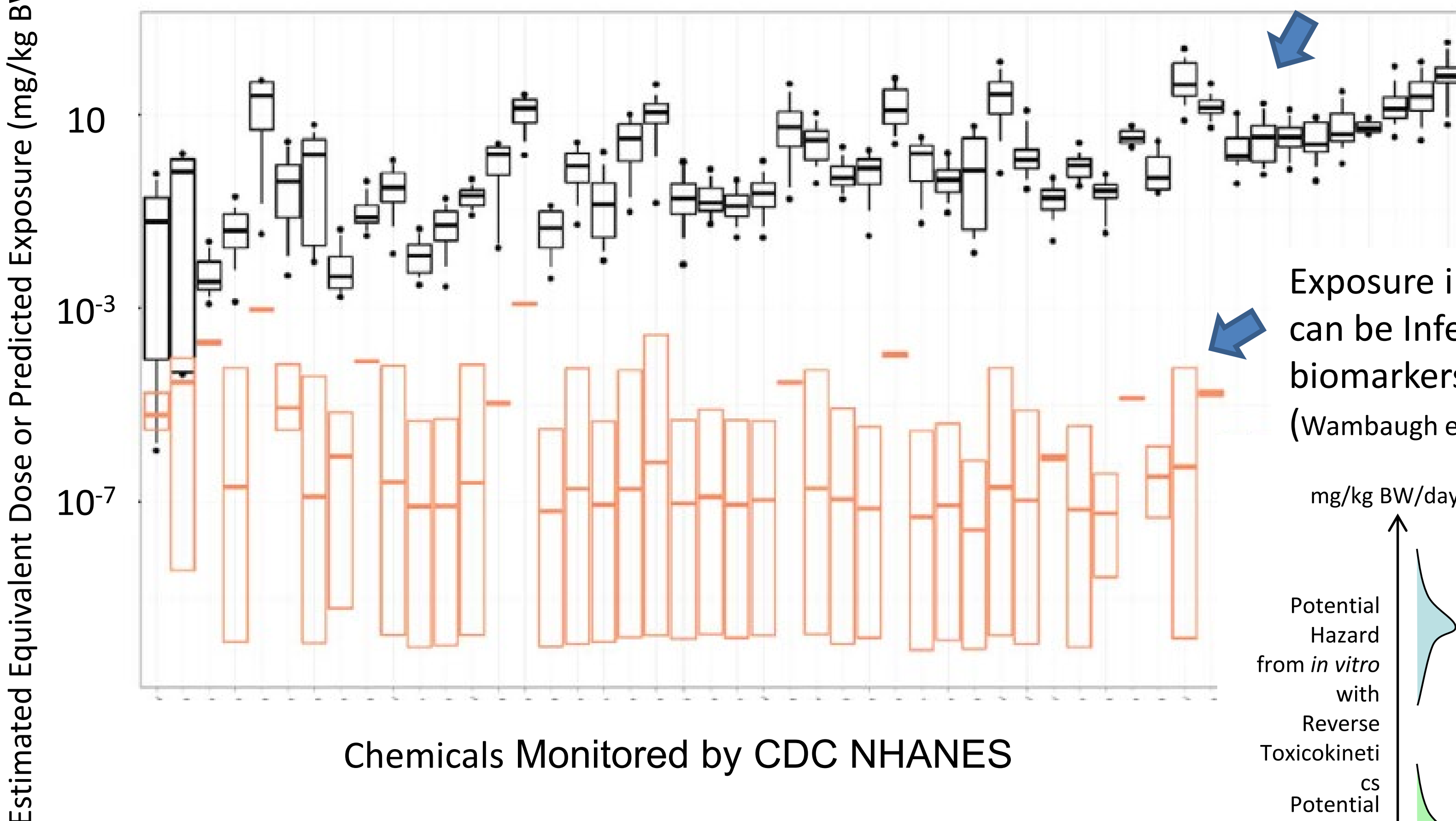
High throughput screening + HTTK-based *in vitro-in vivo* extrapolation (IVIVE) can predict a dose (mg/kg bw/day) that might be adverse



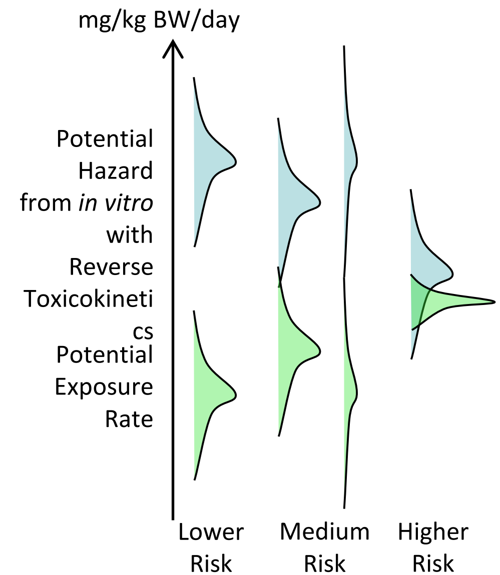
High throughput models exist to make predictions of exposure via specific, important pathways such as residential product use, diet, and environmental fate and transport

High Throughput Risk Prioritization

ToxCast + HTTK can estimate doses needed to cause bioactivity



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



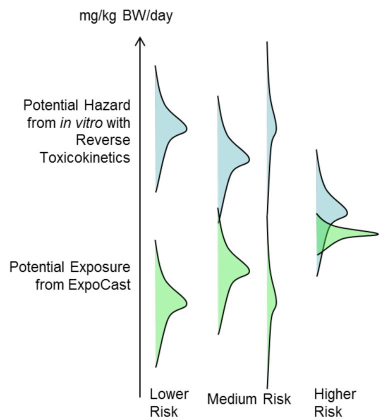
National Health and Nutrition Examination Survey (NHANES) is an ongoing survey that covers ~10,000 people every two years

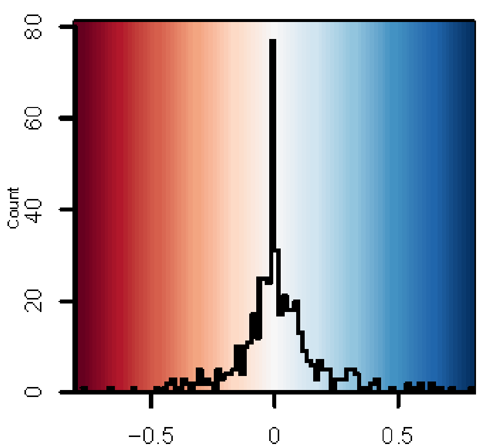
Ring *et al.* (2017)

EPA United States Environmental Protection Agency

Life-stage and Demographic Specific Predictions

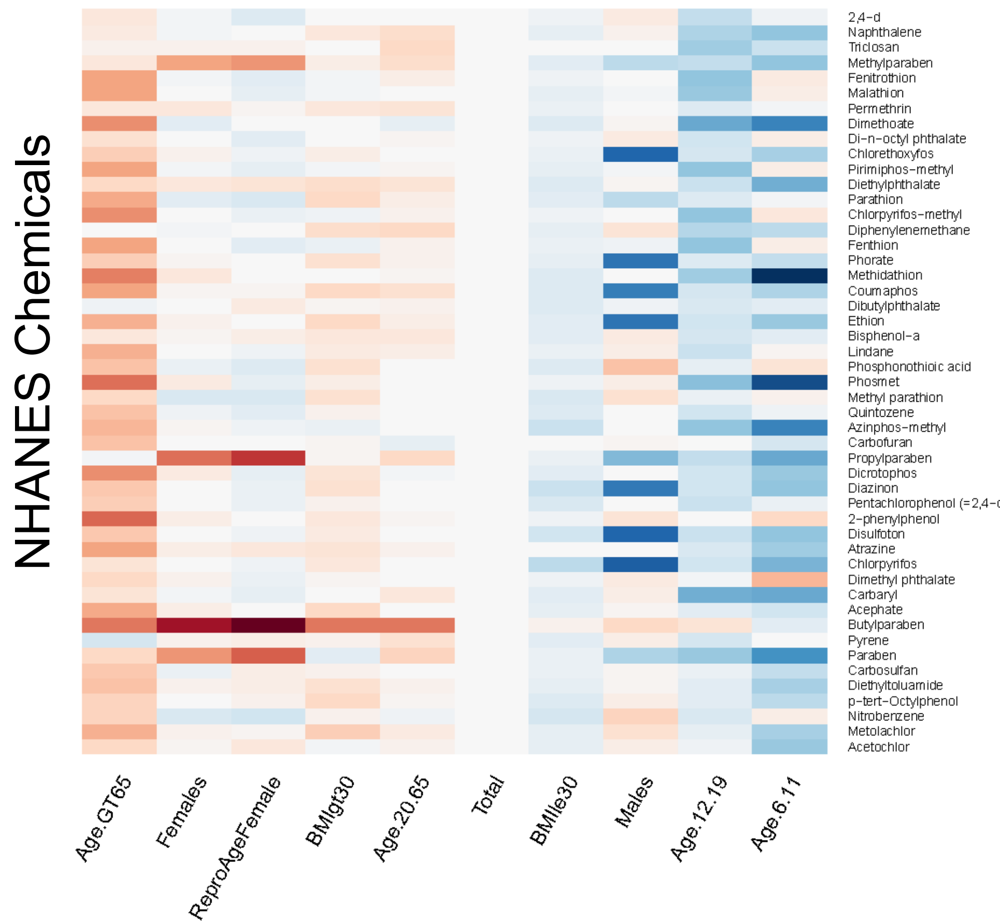
- Can calculate margin between bioactivity and exposure for specific populations





Change in Risk Relative to Total Population

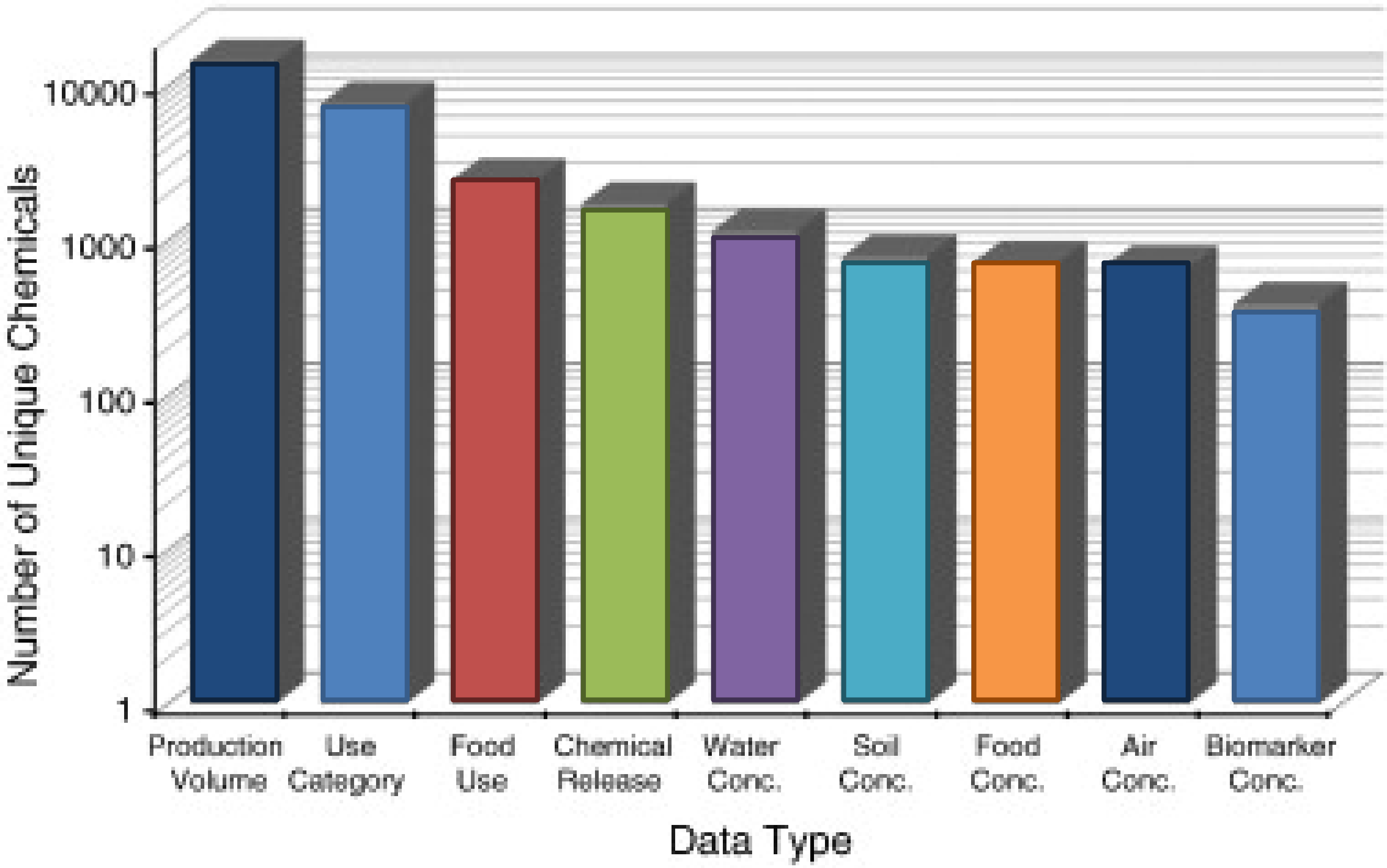
Change in Activity:Exposure Ratio



Ring *et al.* (2017)

Limited Available Data for Exposure Estimation

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



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

What Do We Know About Exposure? Exposure Models

- Human chemical exposures can be coarsely grouped into "**near field**" sources that are close to the exposed individual (consumer or occupational exposures) '**far-field**' scenarios wherein individuals are exposed to chemicals that were released or used far away (ambient exposure) (Arnot *et al.*, 2006).

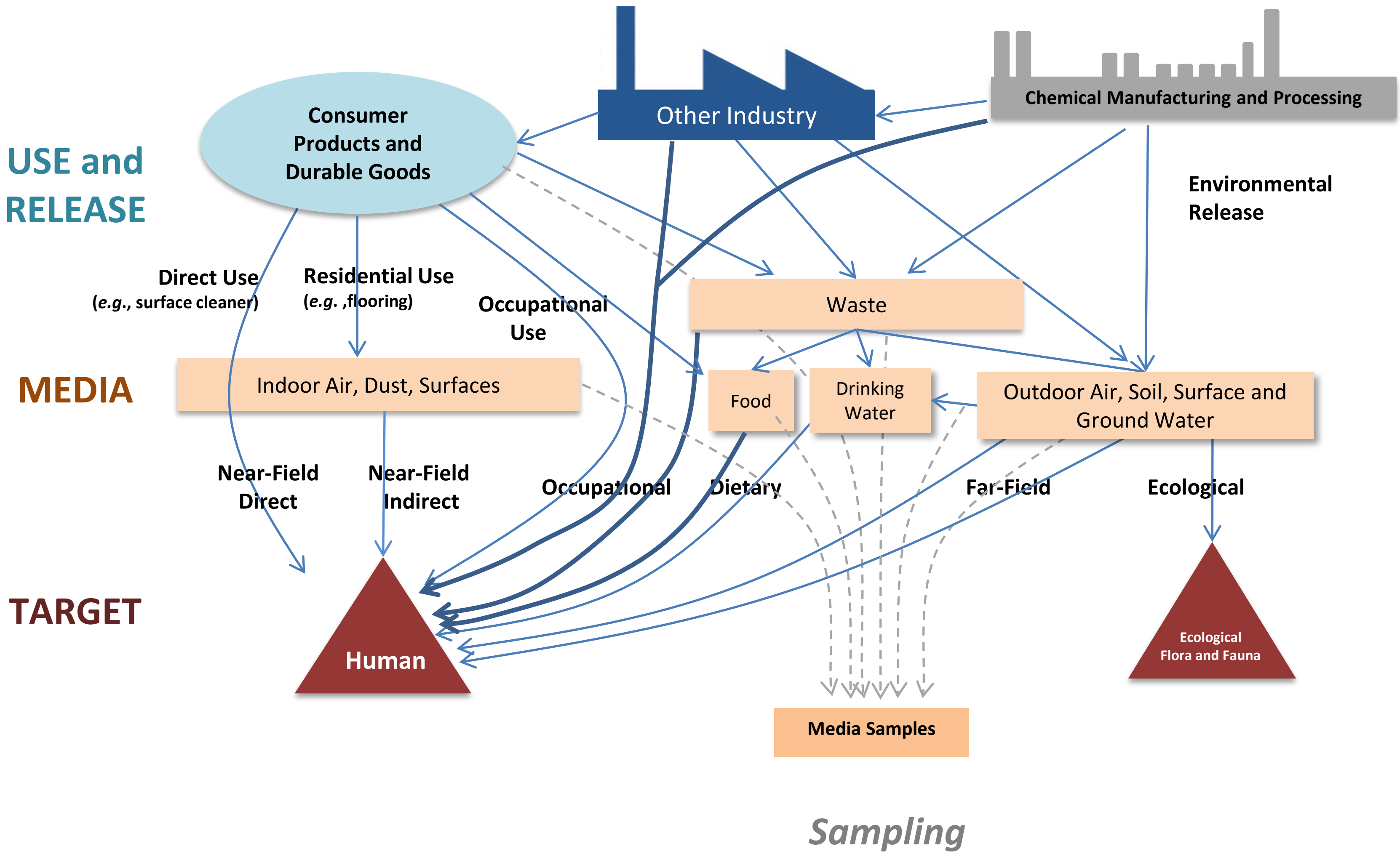
- A model captures knowledge and a hypothesis of how the world works (MacLeod *et al.*, 2010)

- 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

"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

Forecasting Exposure is a Systems Problem



 Office of Research and Development

Figure from Kristin Isaacs

Source-based Exposure Pathways

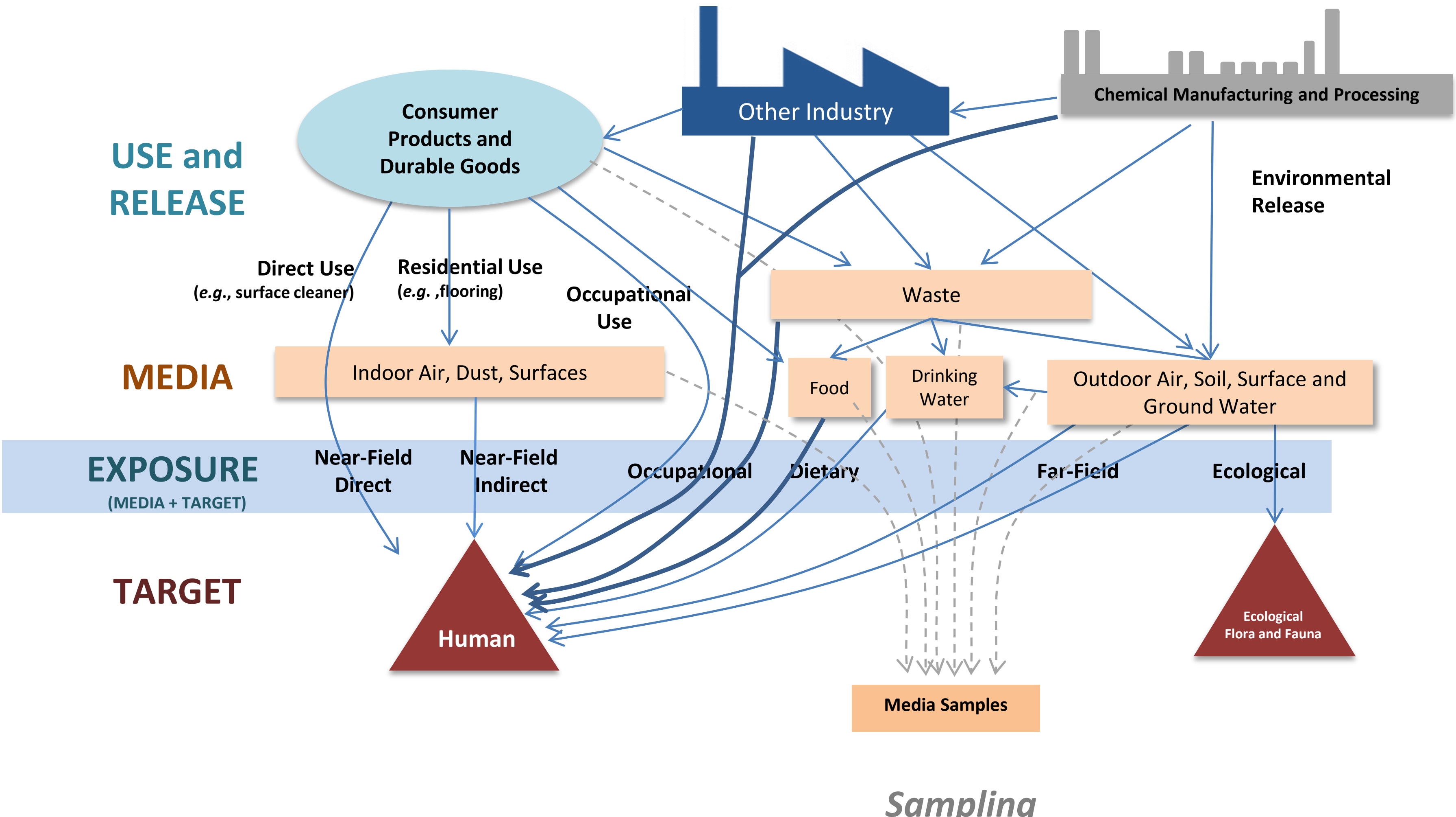


Figure from Kristin Isaacs

United States Environmental Protection Agency

The Exposure Event is Often Unobservable

- The exposure pathway is the actual interaction of the receptor and media, e.g. consuming potato chips

- For humans in particular, these events are often unobserved and for many reasons (including ethics and privacy) may remain unobservable

  - *Did you eat the serving size or the whole bag of potato chips?*

- **Either predict** exposure using data and models up-stream of the exposure event

- **Or infer** exposure pathways from down-stream data, especially biomarkers of exposure

Figure modified from original by Kristin Isaacs

Models to Predict Exposure

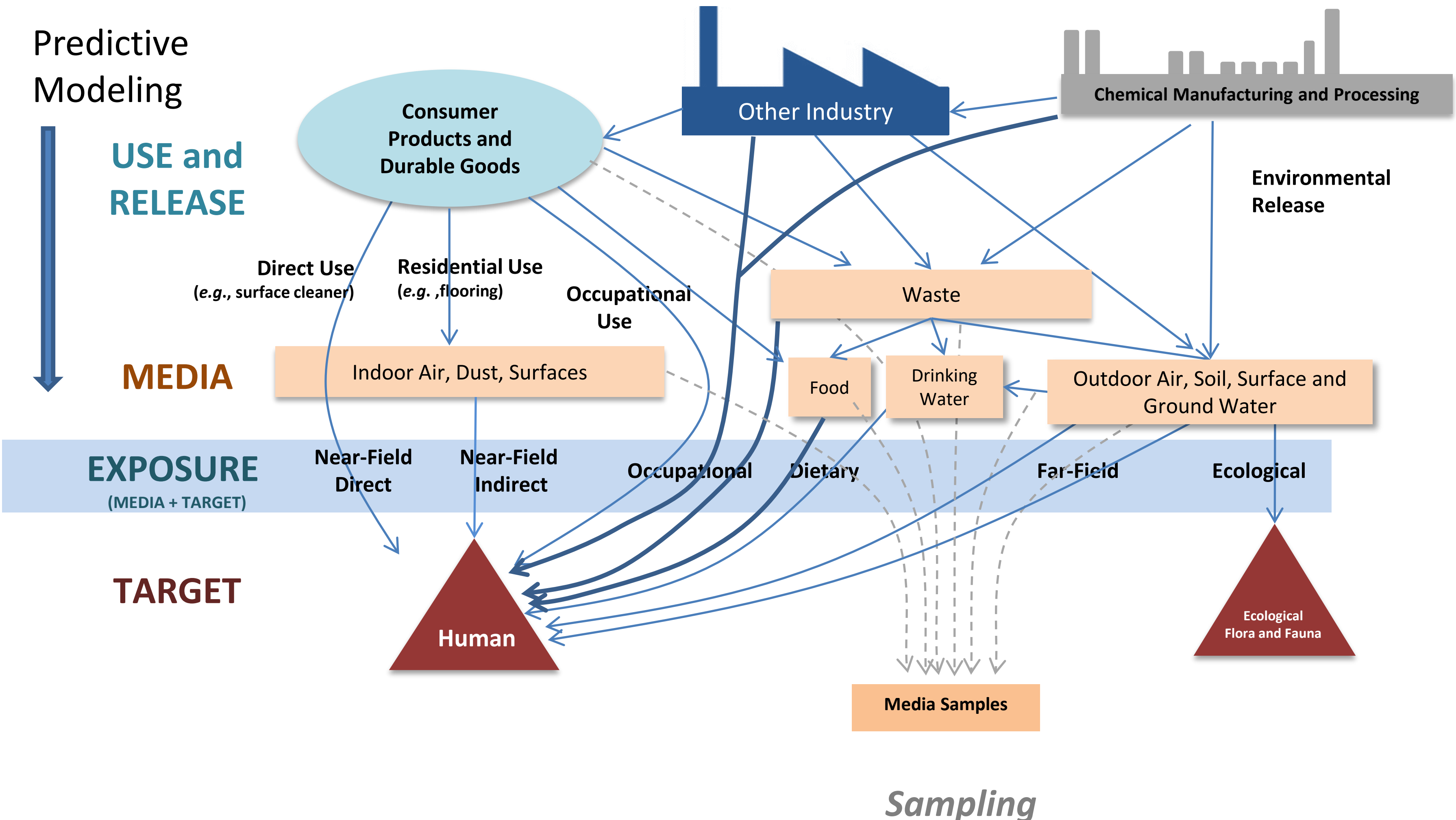


Figure from Kristin Isaacs

What Do We Know About Exposure?

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**

nhanes

National Health and Nutrition Examination Survey

Monitoring Data

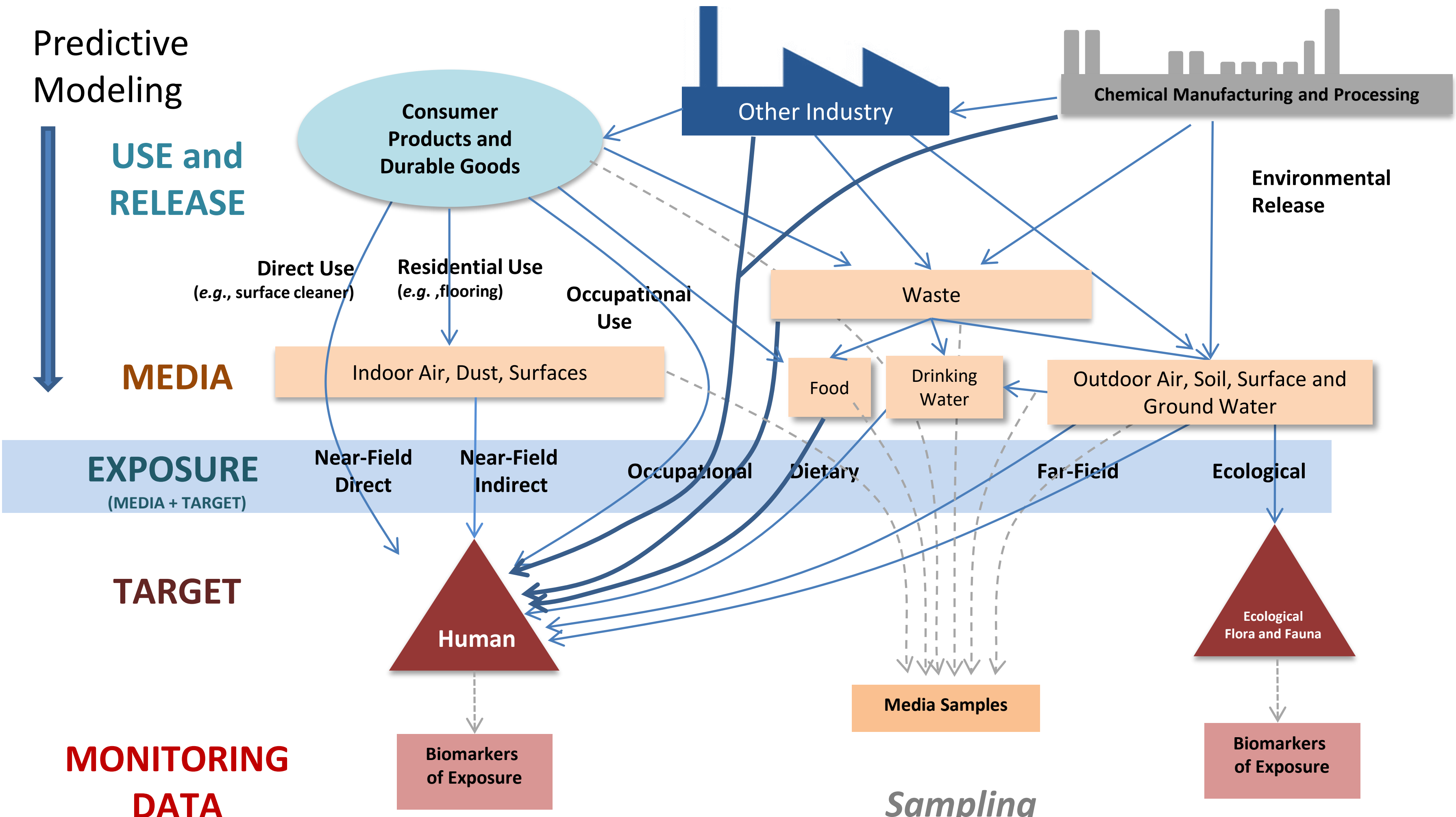


Figure from Kristin Isaacs

EPA United States Environmental Protection Agency

The Six Degrees of Kevin Bacon

On the Solvability of the Six Degrees of Kevin Bacon Game
A Faster Graph Diameter and Radius Computation Method

Michele Borassi[1], Pierluigi Crescenzi[2], Michel Habib[3], Walter Kosters[4], Andrea Marino[5,*], and Frank Takes[4]

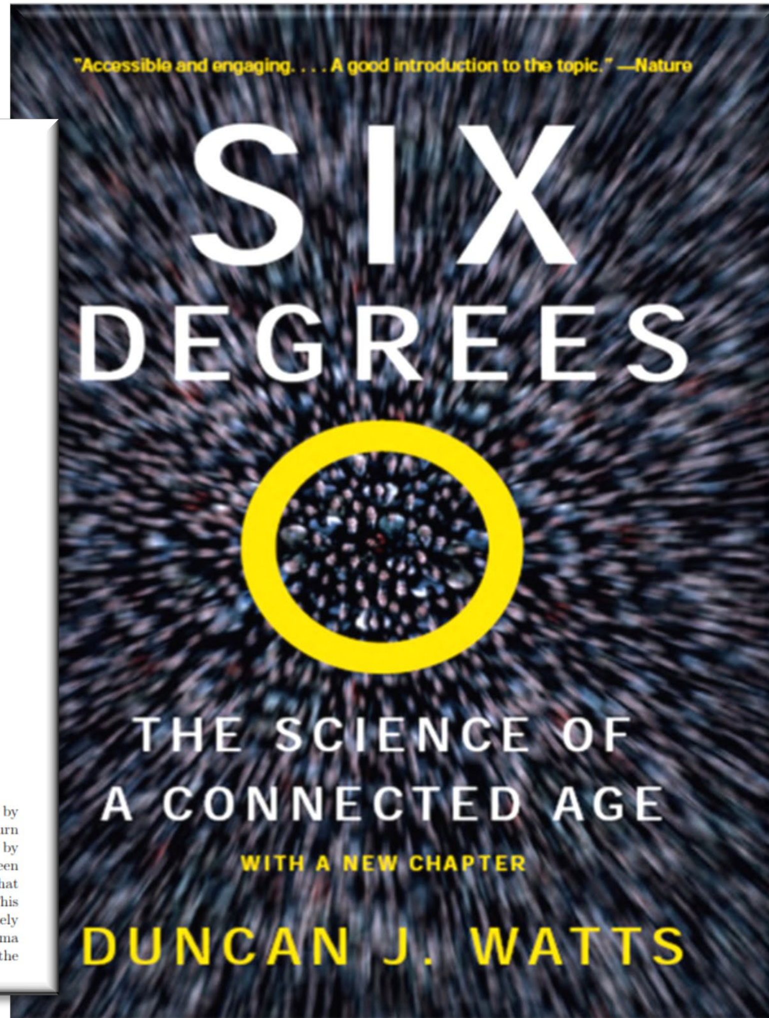
[1] IMT Institute of Advanced Studies, Lucca, Italy
[2] Dipartimento di Sistemi e Informatica, Università di Firenze, Italy
[3] LIAFA, UMR 7089 CNRS & Université Paris Diderot - Paris 7, France
[4] Leiden Institute of Advanced Computer Science, Leiden University, The Netherlands
[5] Dipartimento di Informatica, Università di Milano, Italy

Abstract. In this paper, we will propose a new algorithm that computes the radius and the diameter of a graph $G = (V, E)$, by finding bounds through heuristics and improving them until exact values can be guaranteed. Although the worst-case running time is $\mathcal{O}(|V| \cdot |E|)$, we will experimentally show that, in the case of real-world networks, it performs much better, finding the correct radius and diameter value after 10–100 BFSes instead of $|V|$ BFSes (independent of the value of $|V|$), and thus having running time $\mathcal{O}(|E|)$. Apart from efficiency, compared to other similar methods, the one proposed in this paper has three other advantages. It is more robust (even in the worst cases, the number of BFSes performed is not very high), it is able to simultaneously compute radius and diameter (halving the total running time whenever both values are needed), and it works both on directed and undirected graphs with very few modifications. As an application example, we use our new algorithm in order to determine the solvability over time of the "six degrees of Kevin Bacon" game.

1 Introduction

The six degrees of separation game is a trivia game which has been inspired by the well-known social experiment of Stanley Milgram [11], which was in turn a continuation of the empirical study of the structure of social networks by Michael Gurevich [7]. Indeed, the notion of six degrees of separation has been formulated for the first time by Frigyes Karinthy in 1929, who conjectured that any two individuals can be connected through at most five acquaintances. This conjecture has somehow been experimentally verified by Milgram and extremely popularized by a theater play of John Guare, successively adapted to the cinema by Fred Schepisi. The corresponding game refers to a social network, such as the

* The fifth author was supported by the EU-FET grant NADINE (GA 288956).



KEVIN BACON AND GRAPH THEORY

Brian Hopkins

DRESS: Department of Mathematics, Saint Peter's College, Jersey City NJ 07306 USA. bhopkins@spc.edu.

STRACT: The interconnected world of actors and movies is a familiar, rich example for graph theory. This paper gives the history of the "Kevin Bacon Game" and makes extensive use of a Web site to analyze the underlying graph. The main content is the classroom development of the weighted average to determine the best choice of "center" for the graph. The article concludes with additional student activities and some responses to the material.

YWORDS: Cinema, finite mathematics, graph theory, popular culture, six degrees of separation, weighted averages.

1 INTRODUCTION

ph theory is the mathematics of connections. It has wide applications to , interconnected systems: transportation networks, epidemiology, and Internet, to name just a few. But we teach graph theory with pictures handful of dots and lines. There is one large system that is easy to work , thanks to a Web site run by the University of Virginia, Department omputer Science. The Oracle of Bacon at Virginia [6] uses the Internet ie Database [3], which documents almost all of cinematic history. This is od tool for illustrating complete subgraphs, connected components, and distance between vertices. There is also a nice application of weighted ages. I have used this material in freshman finite mathematics classes mathematics major courses that cover graph theory; students always ond enthusiastically.

5

United States Environmental Protection Agency

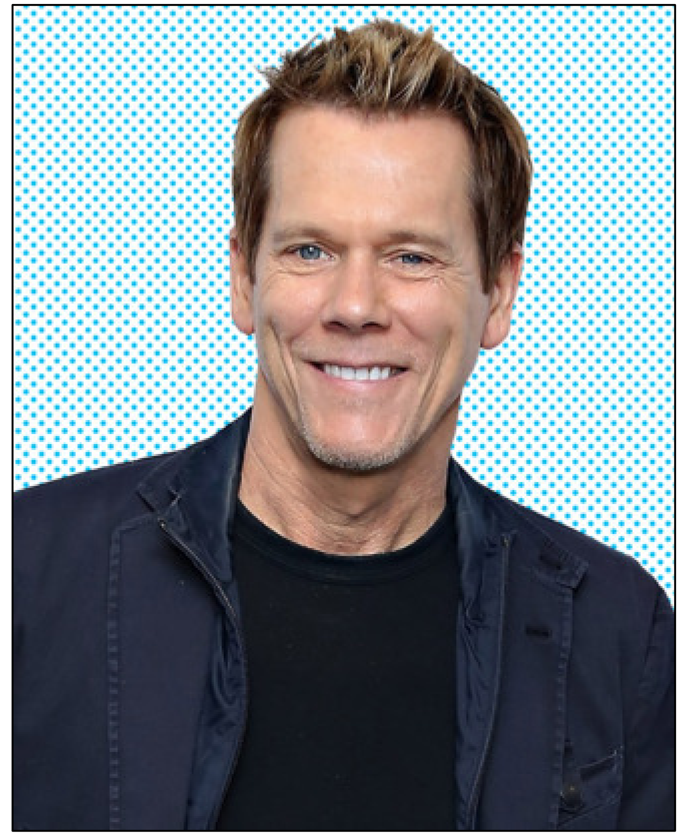
Kevin Bacon

1978

1984

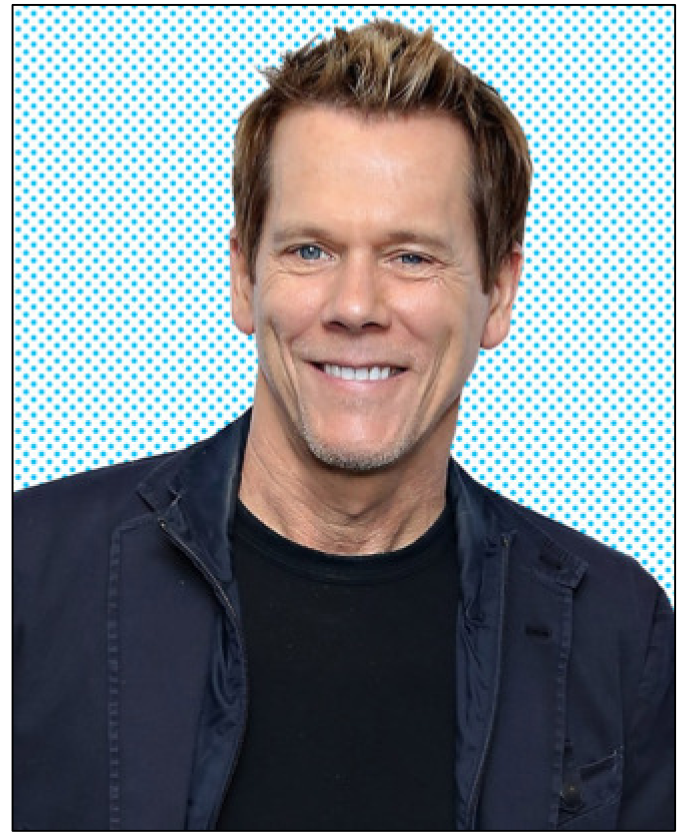
1992

1995



Kevin Bacon

1990



Michael B. Jordan

United States
Environmental Protection
Agency

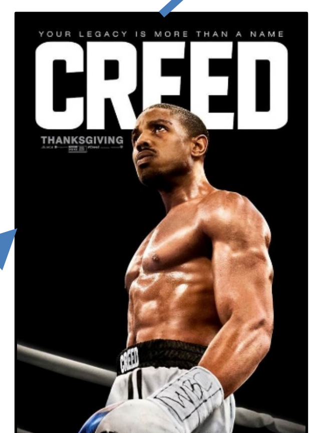
Connectedness to Michael B. Jordan

Frances McDormand
Best Actress Winner 2018

Hail Caesar
McDormand &
Channing Tatum

GI Joe: Retaliation
Tatum & Bruce Willis

Expendables
Willis &
Sylvester Stallone



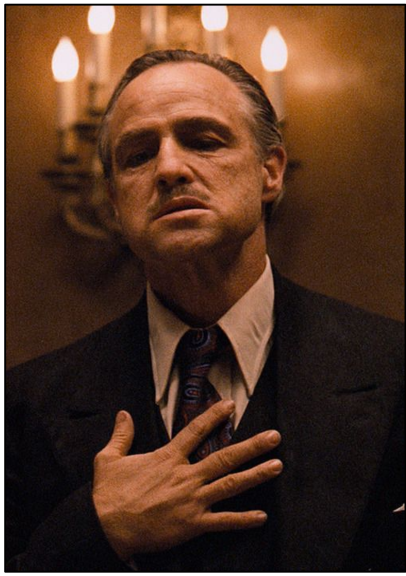
Creed
Stallone & Jordan

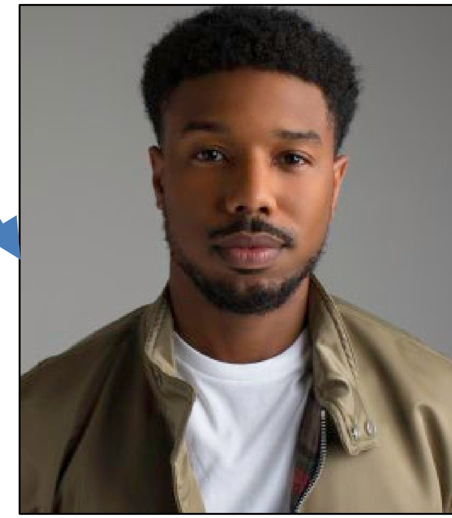
United States Environmental Protection Agency

Connectedness to Michael B. Jordan

Avengers: Infinity War
Paltrow & Chadwick Boseman

Black Panther
Boseman & Jordan





Marlon Brando
Best Actor 1954 and 1972
Died 2004

Superman
with Gene Hackman

The Royal Tenenbaums
Hackman & Gwyneth Paltrow

Small World Networks

United States Environmental Protection Agency

Watts and Strogatz (1998)

letters to nature

Collective dynamics of 'small-world' networks

Duncan J. Watts* & Steven H. Strogatz

Department of Theoretical and Applied Mechanics, Kimball Hall, Cornell University, Ithaca, New York 14853, USA

Networks of coupled dynamical systems have been used to model biological oscillators[1-4], Josephson junction arrays[5,6], excitable media[7], neural networks[8-10], spatial games[11], genetic control networks[12] and many other self-organizing systems. Ordinarily, the connection topology is assumed to be either completely regular or completely random. But many biological, technological and social networks lie somewhere between these two extremes. Here we explore simple models of networks that can be tuned through this middle ground: regular networks 'rewired' to introduce increasing amounts of disorder. We find that these systems can be highly clustered, like regular lattices, yet have small characteristic path lengths, like random graphs. We call them 'small-world' networks, by analogy with the small-world phenomenon[13,14] (popularly known as six degrees of separation[15]). The neural network of the worm *Caenorhabditis elegans*, the power grid of the western United States, and the collaboration graph of film actors are shown to be small-world networks. Models of dynamical systems with small-world coupling display enhanced signal-propagation speed, computational power, and synchronizability. In particular, infectious diseases spread more easily in small-world networks than in regular lattices.

Travers and Milgram (1977):

296 arbitrary individuals in Nebraska and Boston were asked to give a letter to an acquaintance most likely to help it reach a target person in Massachusetts. 64 reached the target person, average number of intermediaries was 5.2

Collins and Chow (1998)

news and views

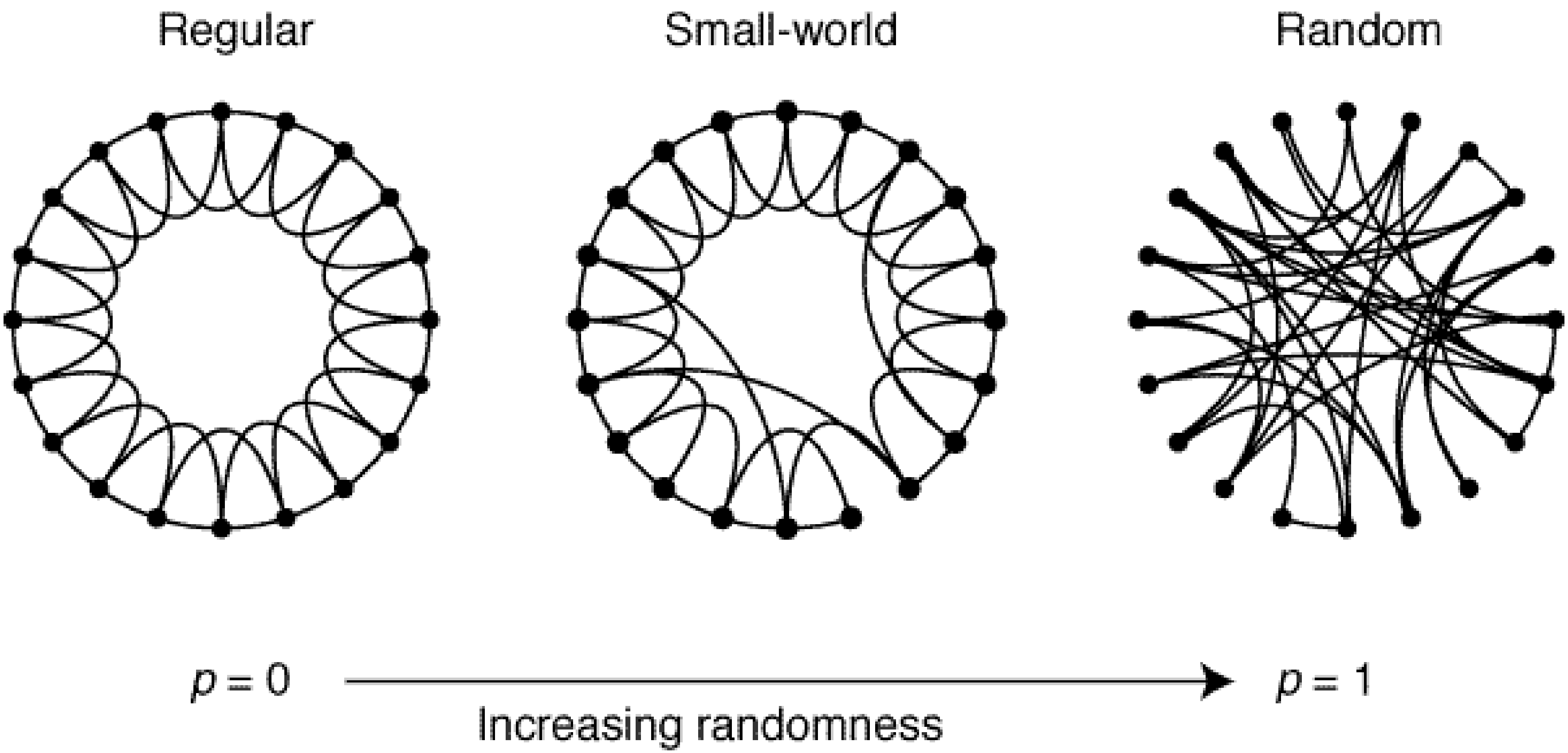
It's a small world

James J. Collins and Carson C. Chow

The concept of Six Degrees of Separation has been formalized in so-called 'small-world networks'. The principles involved could be of use in settings as diverse as improving networks of cellular phones and understanding the spread of infections.

Figure 1 Three degrees. Because Kevin Bacon has appeared in many films, most actors have low Bacon numbers and the game Six Degrees of Kevin Bacon has declined in popularity. It is possible to centre the game around a newer star such as Leonardo DiCaprio. These film stills, running clockwise, show that in this case there are at most three degrees of separation between DiCaprio and Helena Bonham-Carter, through Kate Winslet (*Titanic*, Columbia TriStar; *Sense and Sensibility*, Columbia TriStar), Emma Thompson (*Sense and Sensibility*; *Much Ado About Nothing*, Entertainment Films) and Kenneth Branagh (*Much Ado About Nothing*; *Frankenstein*; Columbia TriStar). Short cuts between cliques could be created in this game through some of DiCaprio's well-connected co-stars such as Sharon Stone (*The Quick and the Dead*; TriStar; not shown).

Complex is Not the Same as Random



Watts and Strogatz (1998)

Models to Infer Exposure

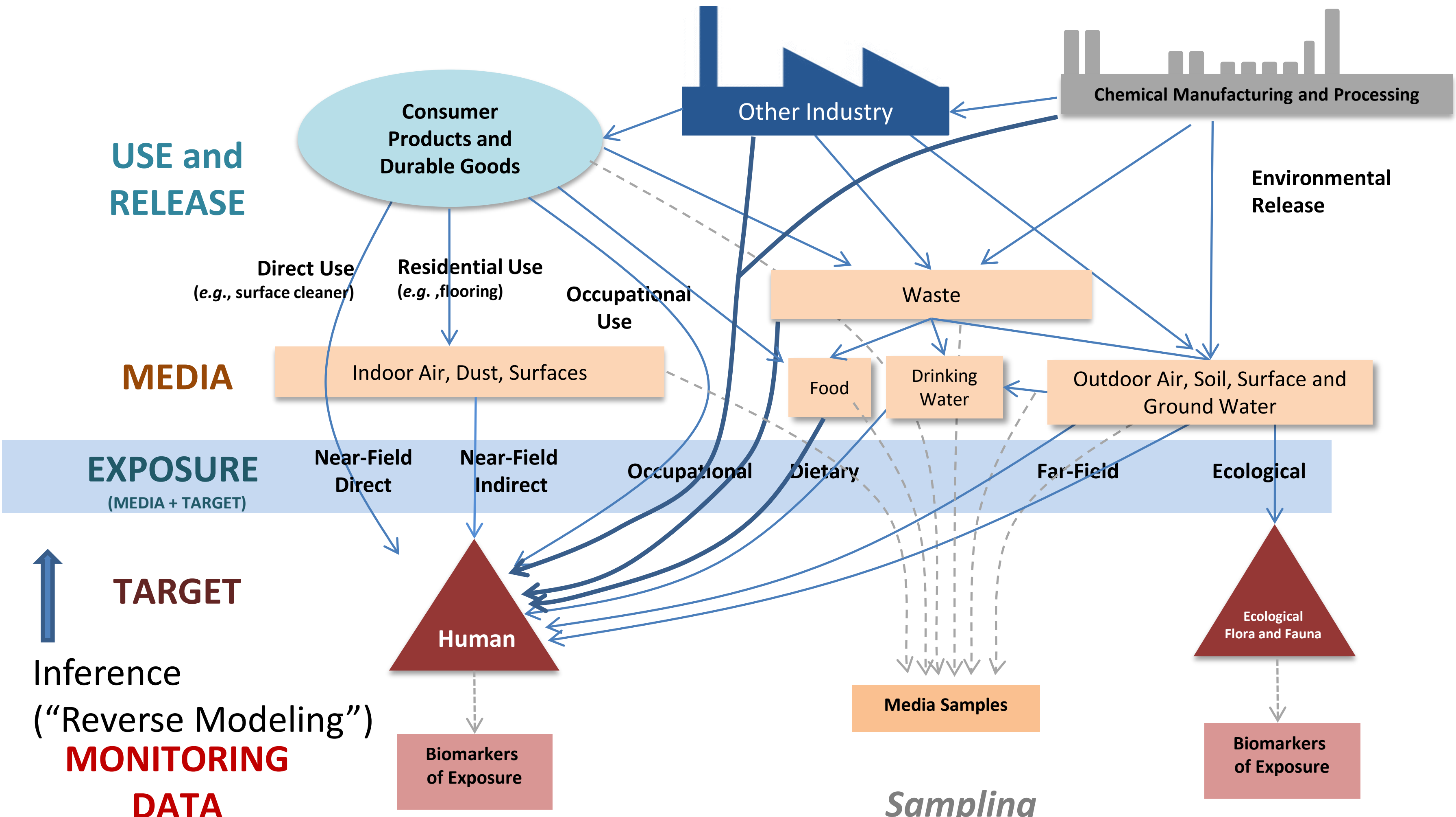
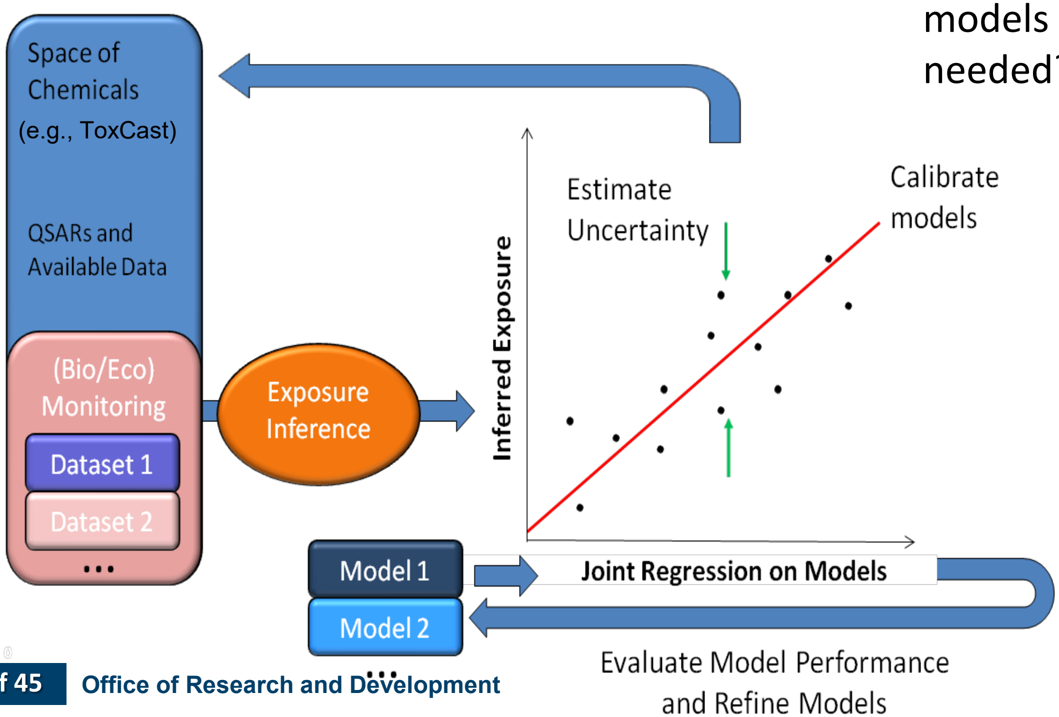


Figure from Kristin Isaacs

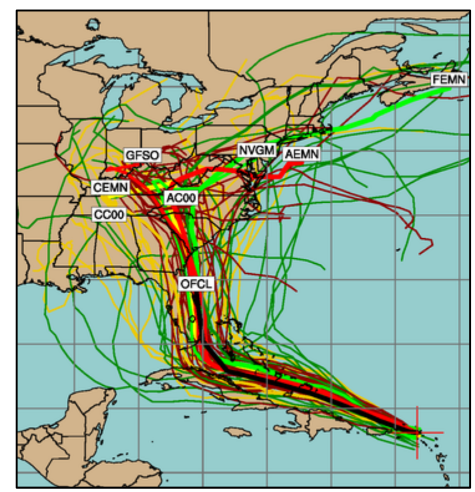
Consensus Exposure Predictions with the SEEM Framework

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

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



Hurricane Path Prediction is an Example of Integrating Multiple Models



Evaluating Models with Monitoring Data

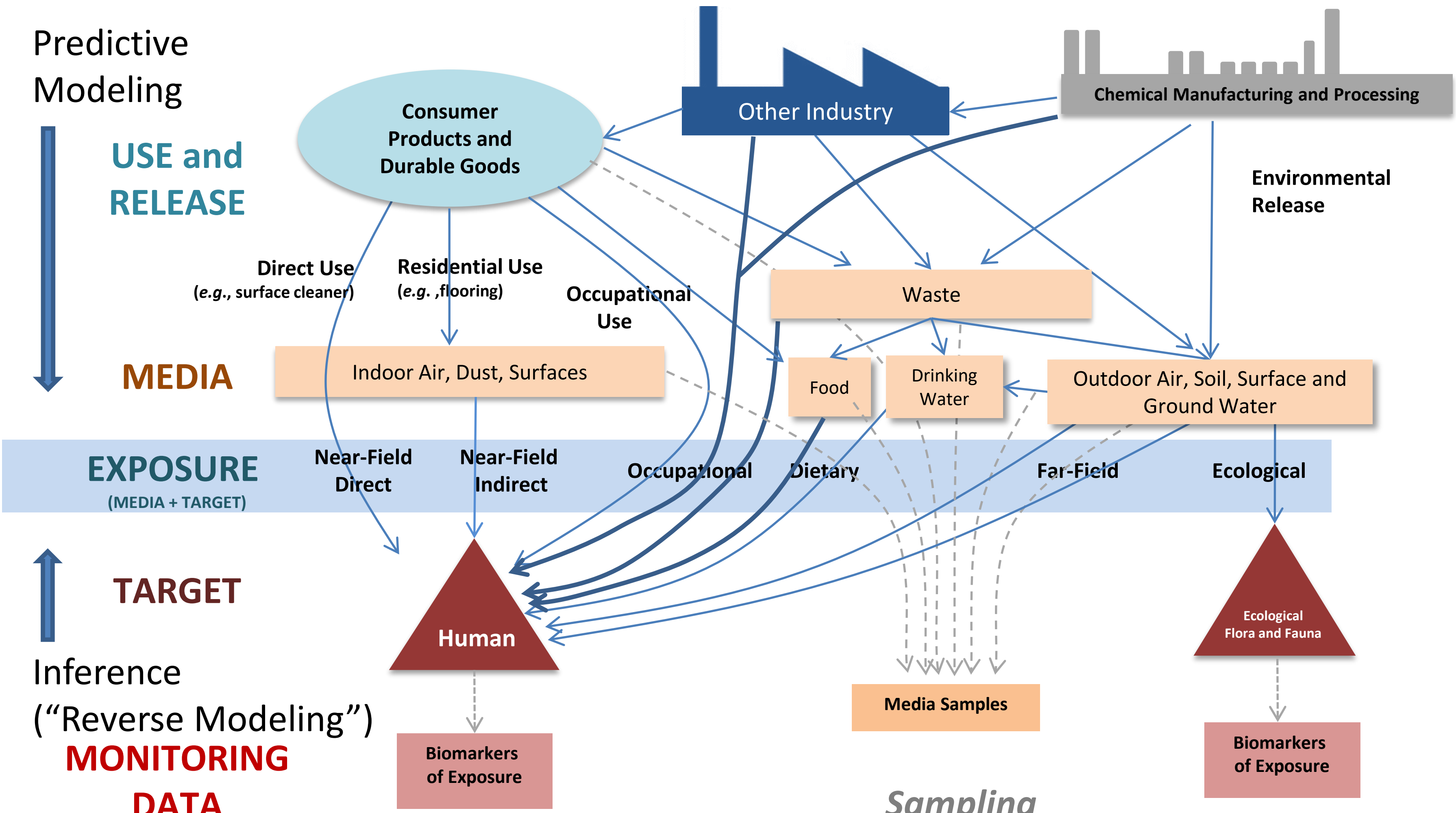
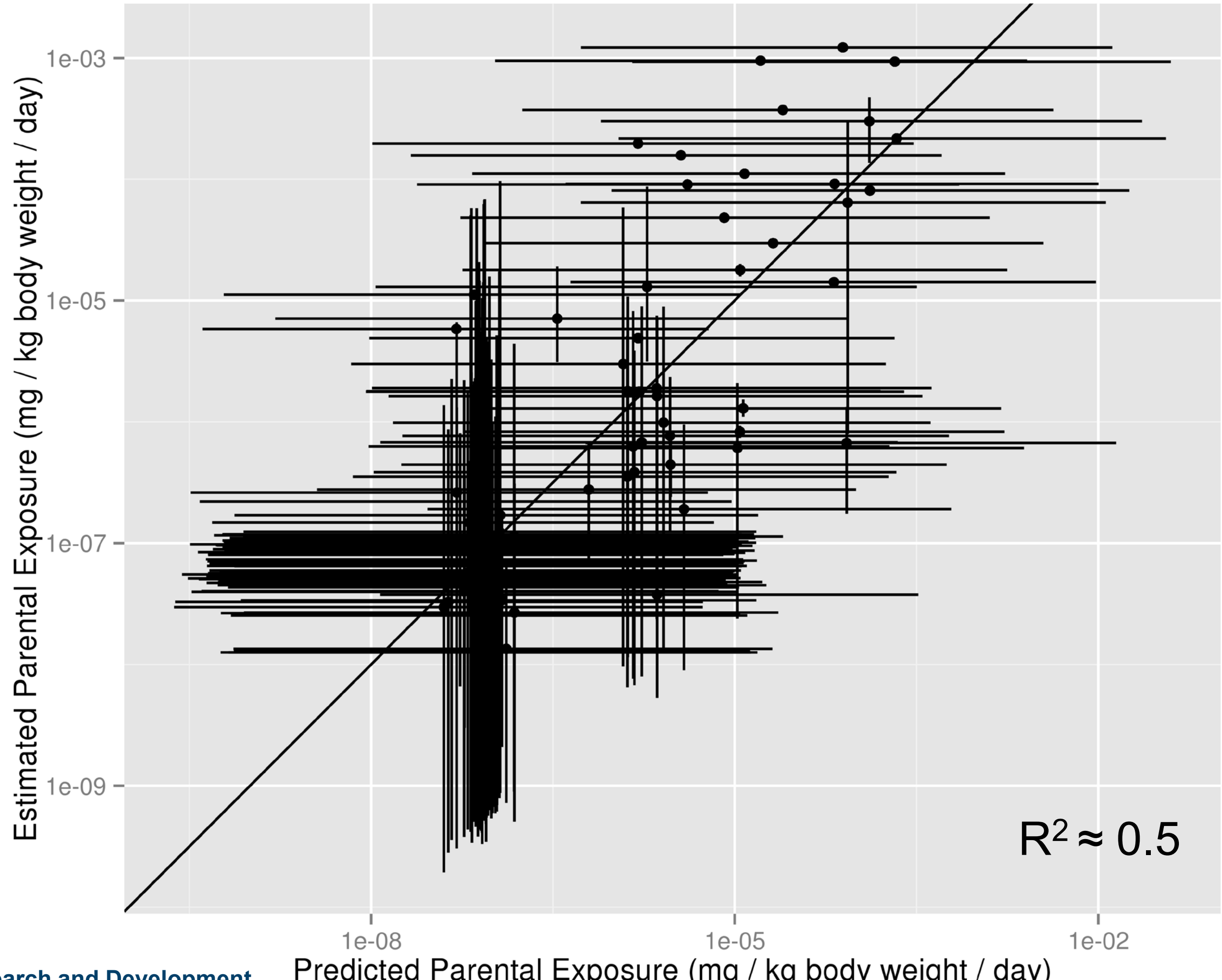


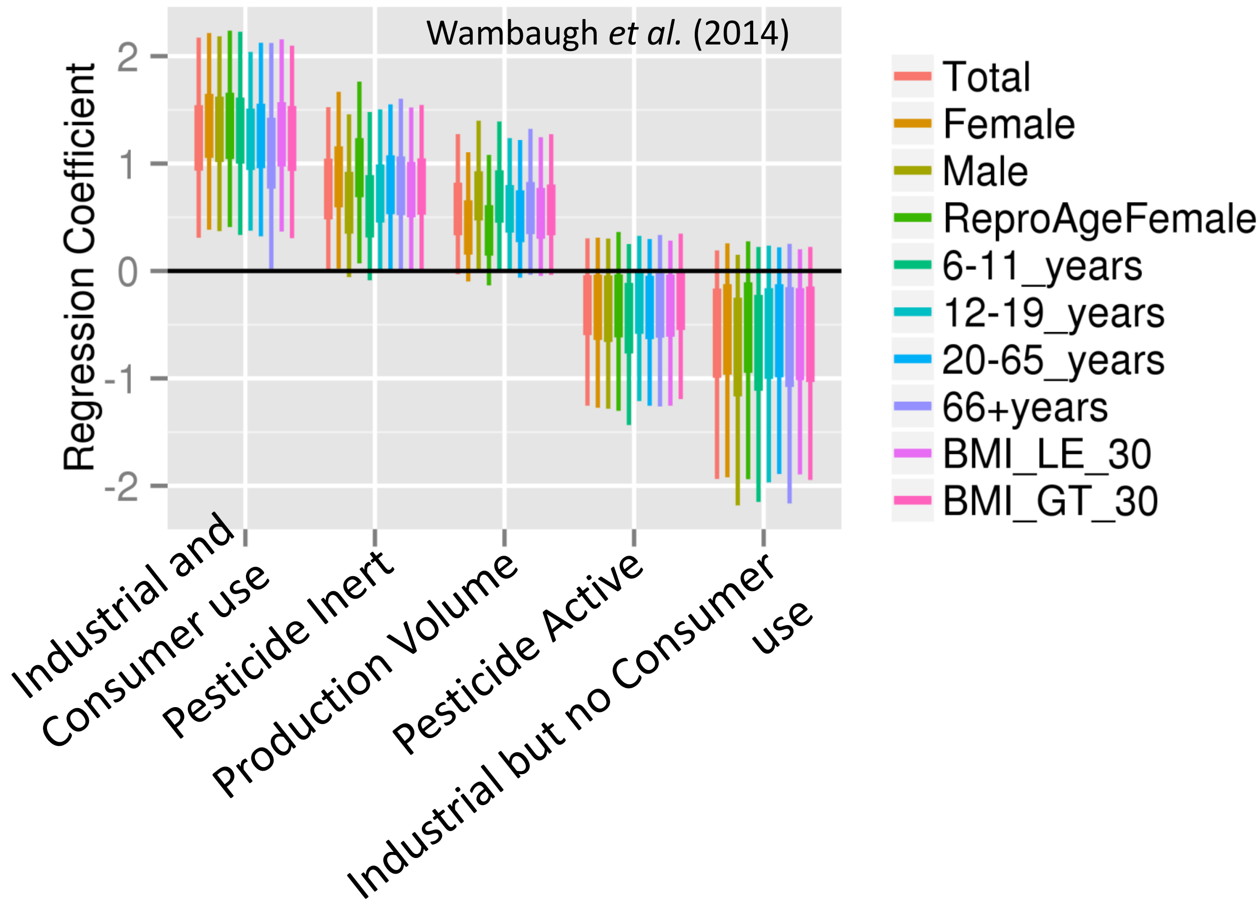
Figure from Kristin Isaacs

SEEM Analysis (circa 2014)



Wambaugh et al, 2014

Heuristics of Exposure



- Five descriptors explain roughly 50% of the chemical-to-chemical variability in median NHANES exposure rates
- Same five predictors work for all NHANES demographic groups analyzed – stratified by age, sex, and body-mass index
- Chemical use identifies relevant pathways
- Some pathways have much higher average exposures (Wallace et al., 1987)

CPCPdb: Material Safety Data Sheets

United States Environmental Protection Agency

Goldsmith et al. (2014):
- ~20,000 product-specific Material Safety Data Sheets (MSDS) curated
- ~2,400 chemicals

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

XXXXXX

XXXXXXXXXX
XXXXXXXXXX

Material Safety Data Sheet

COM-35604

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

II Health Hazard Data

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

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

III Hazardous Ingredients

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

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

IV Special Protection and Precautions

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

Avoid contact with food.

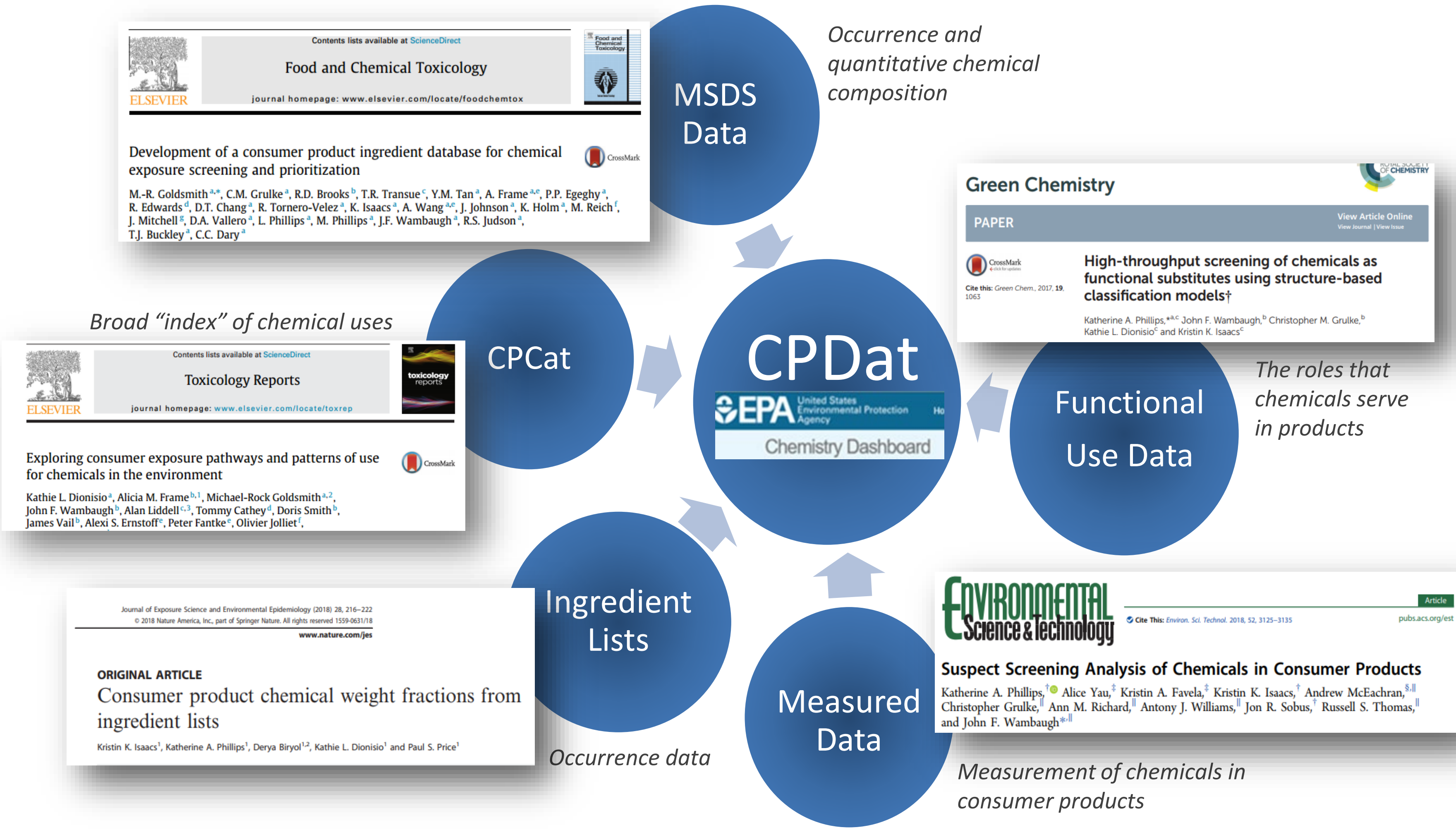
V Transportation and Regulatory Data

U.S. DOT Hazard Class: Not restricted

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

EPA CERCLA/SARA TITLE III:

What Do We Know About Chemical Use? Chemicals and Products Database



https://comptox.epa.gov/dashboard

Slide from Kristin Isaacs

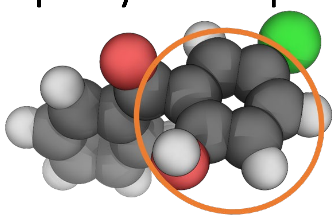
Predicting Function Based on Structure

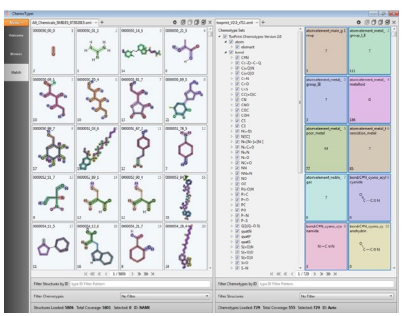
Machine Learning Based Classification Models

(Random Forest, Breiman, 2001)

Use Database (FUSE)

Chemical Structure and Property Descriptors





Prediction of Of Potential Alternatives from Chemical Libraries

Phillips et al. (2017)

Collaboration on High Throughput Exposure Predictions

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

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

Ring et al., 2019

Predicting Exposure Pathways

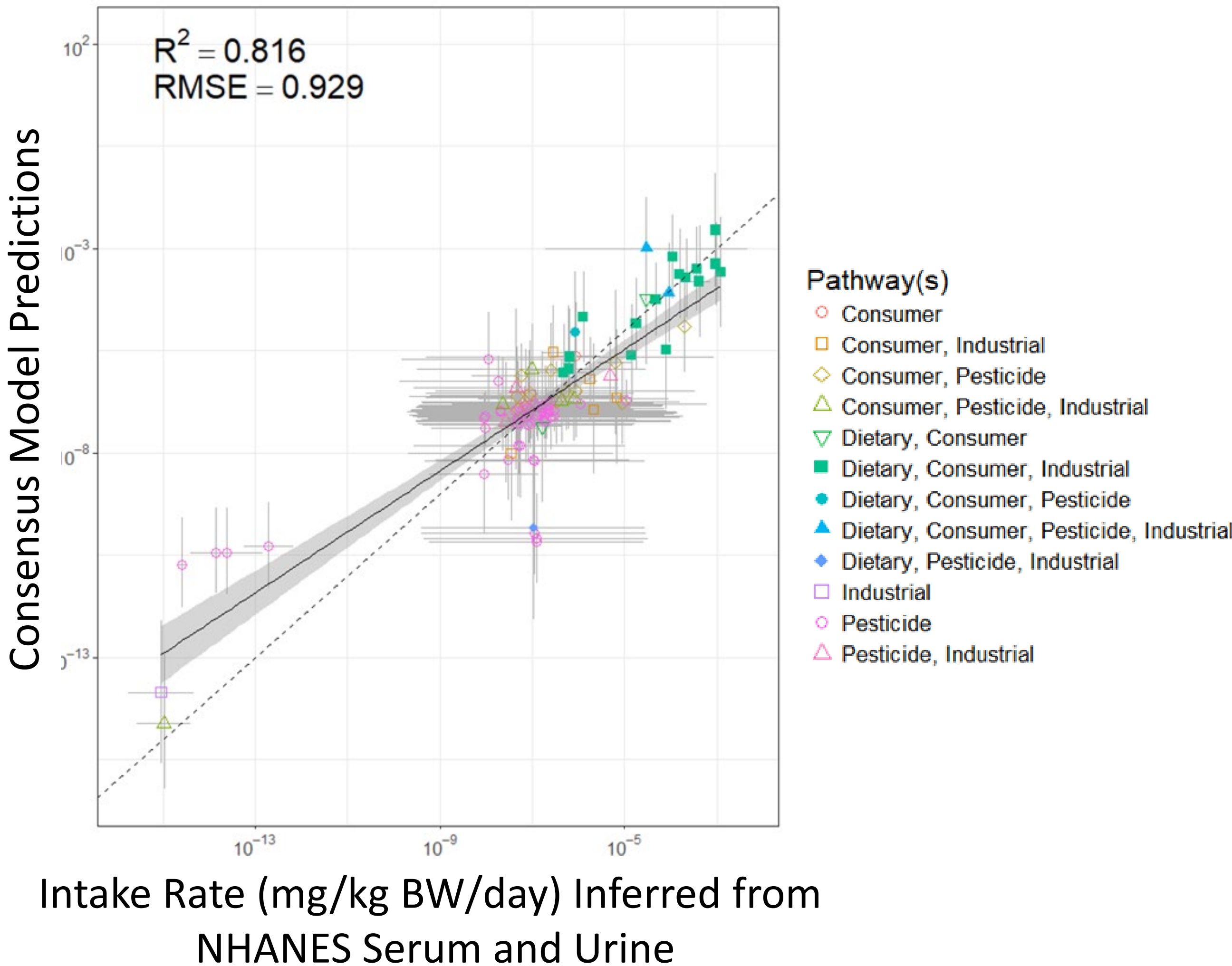
We use the method of Random Forests to relate chemical structure and properties to exposure pathway

	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., 2019

United States Environmental Protection Agency

- Machine learning models were built for each of four exposure pathways
- Pathway predictions can be used for large chemical libraries
- Use prediction (and accuracy of prediction) as a prior for Bayesian analysis
- Each chemical may have exposure by multiple pathways



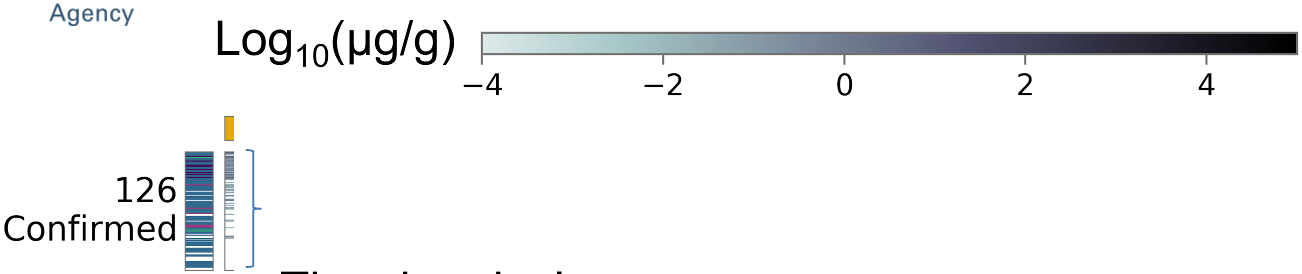
Ring et al., 2019

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

- Not everything is required to have MSDS sheets
- Models present one way forward, but data is always preferable
- New analytic techniques may also allow insight in to the chemical composition of diverse environmental media including household products
- 100 household products from a major U.S. retailer were analyzed, tentatively identifying 1,632 chemicals, 1,445 which were not in EPA's database of consumer product chemicals (Phillips *et al., ES&T just accepted*)

"I'm searching for my keys."

Measuring Chemicals in Household Items

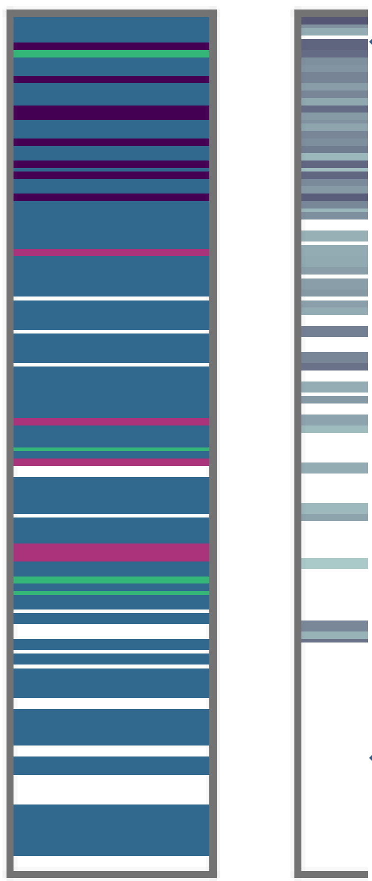


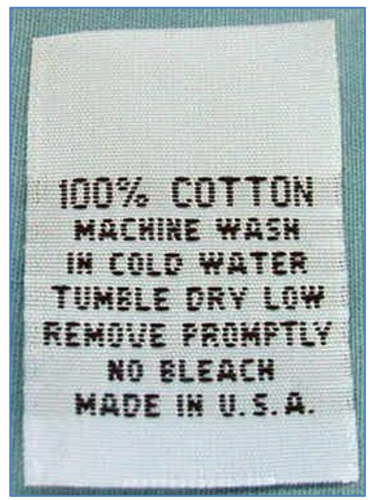
Chemical Category
- ToxCast
- Potent ER
- Flame Retardant
- Chemical in ⩾ 25 Products

The chemicals found in a cotton shirt

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

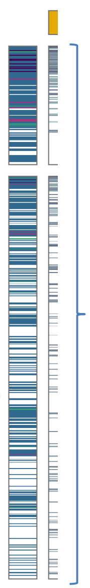
Measuring Chemicals in Household Items



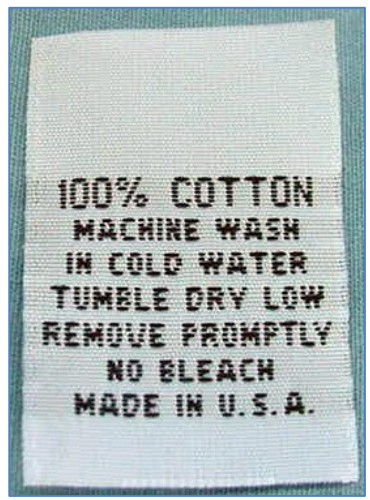


Phillips et al. (*ES&T just accepted*)

Measuring Chemicals in Household Items



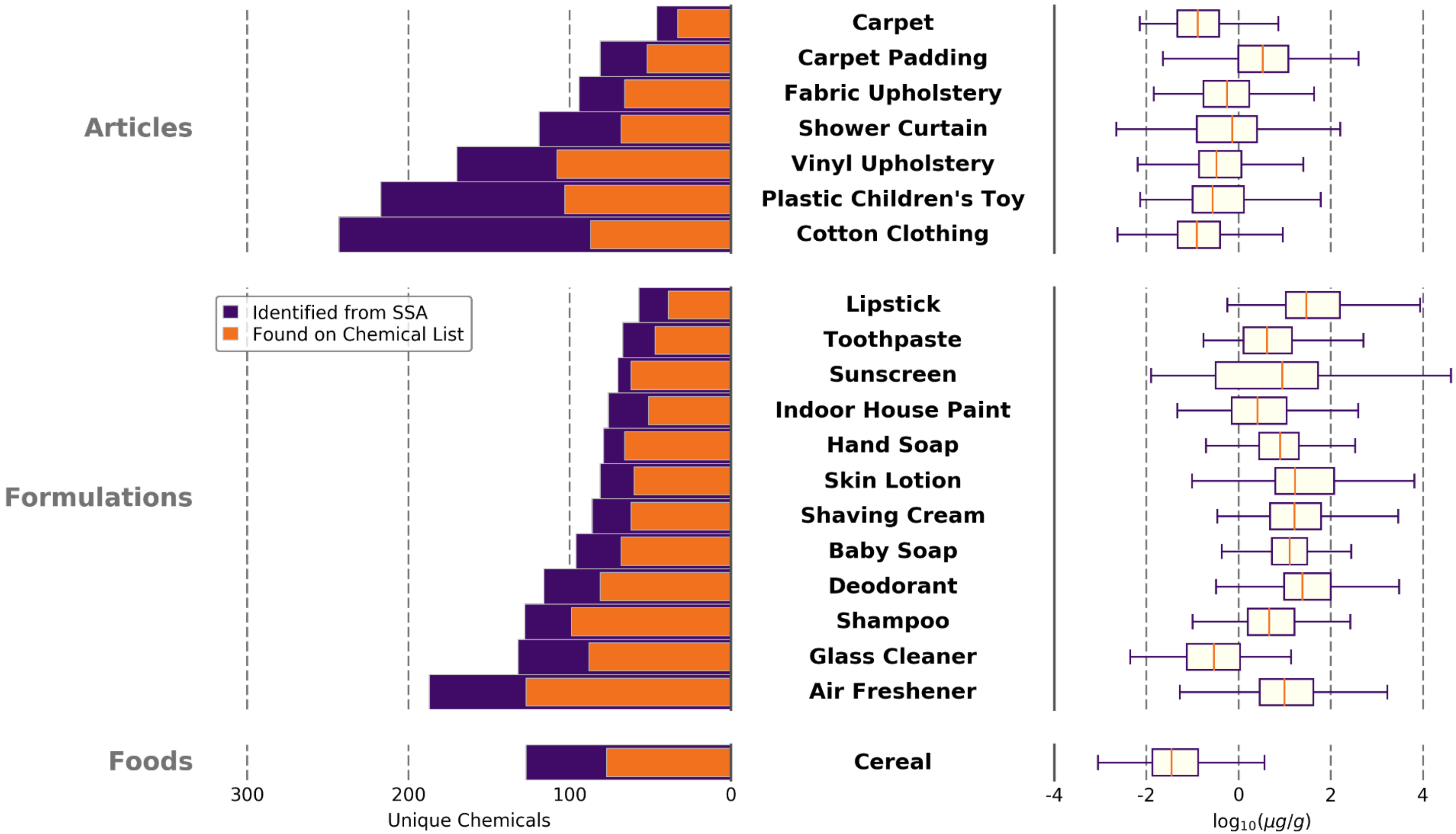
The chemicals found in a cotton shirt



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

Product Scan Summary

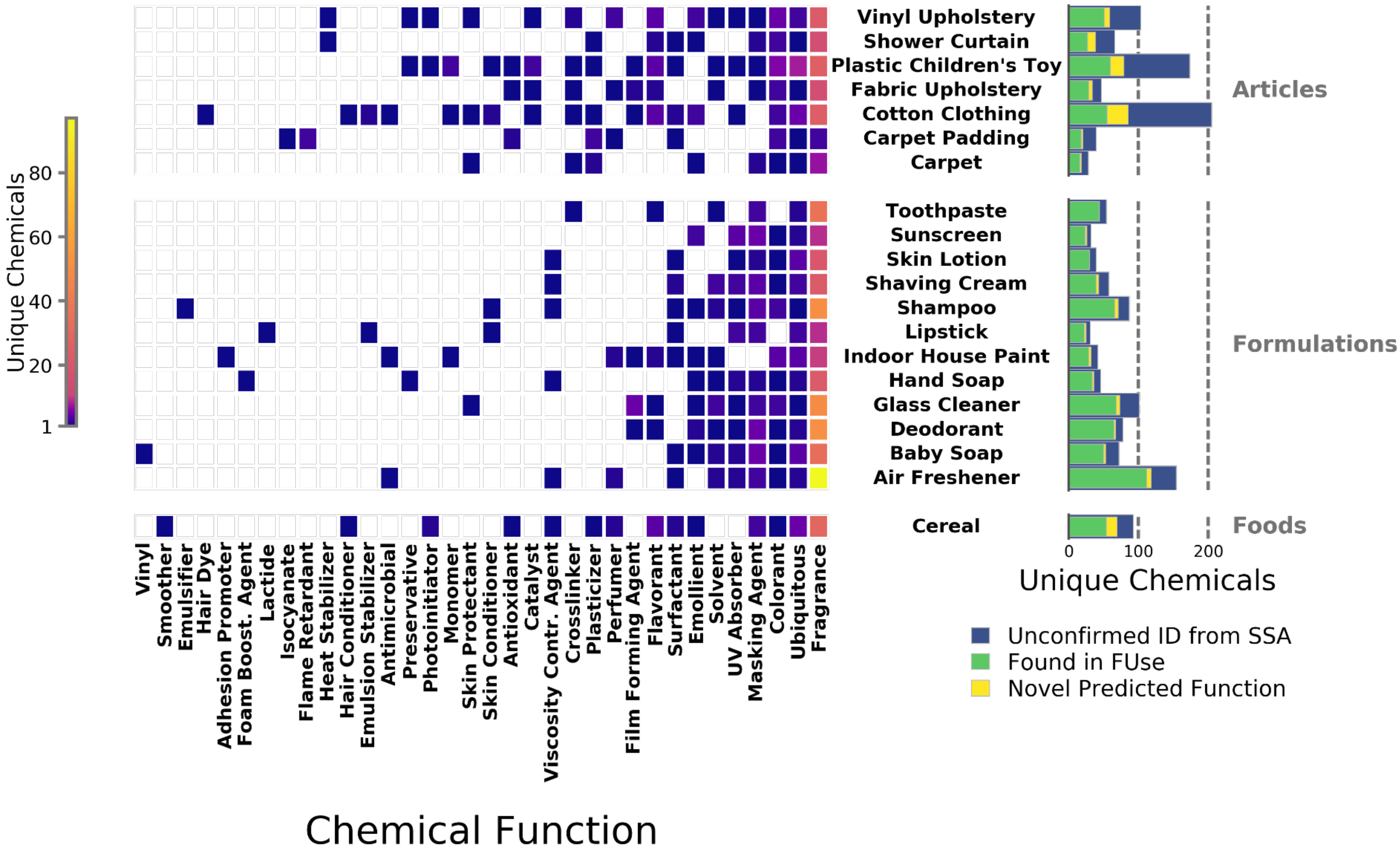
Of 1,632 chemicals confirmed or tentatively identified, 1,445 were not present in CPCPdb (Goldsmith, et al., 2015)



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

Predicting Chemical Function

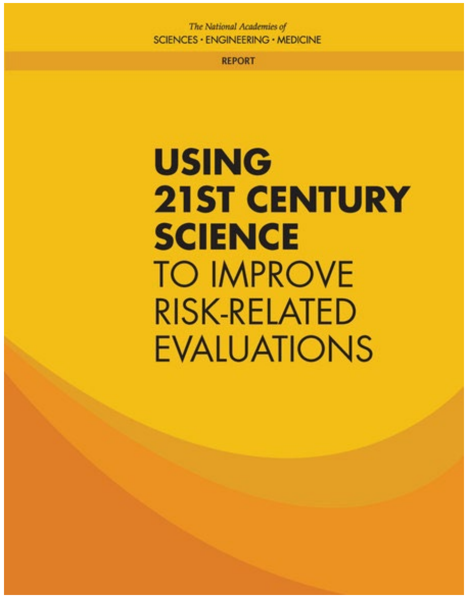
Using the methods of Phillips *et al*., (2017):

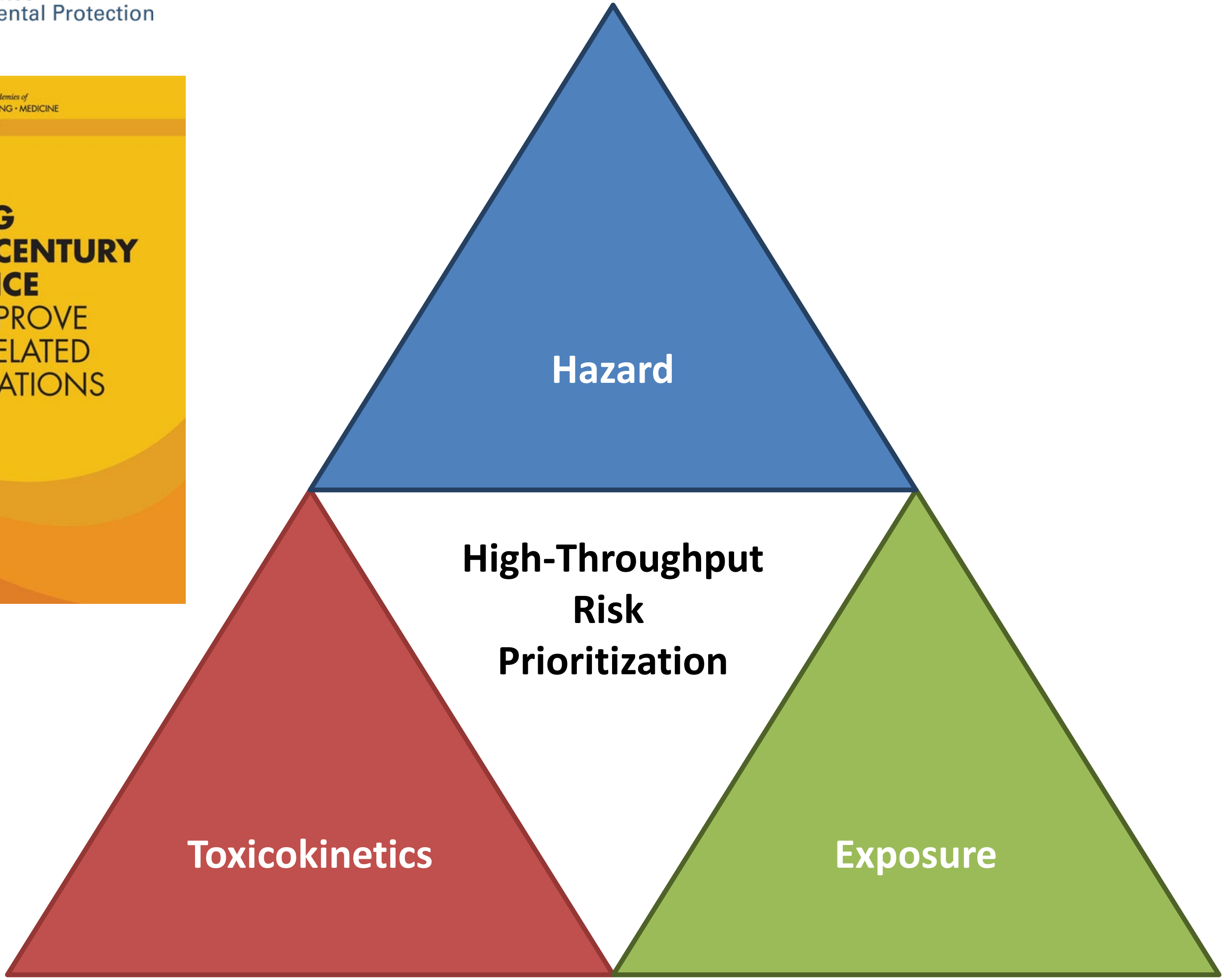


Phillips et al. (*ES&T just accepted*)

Exposure-Based Priority Setting

United States
Environmental Protection
Agency





United States Environmental Protection Agency

Conclusions

- We would like to know more about the risk posed by thousands of chemicals in the environment – which ones should we start with?
  - High throughput screening (HTS) provides one path forward for identifying potential hazard, but the real world is complicated by toxicokinetics, mixtures, variability (and more)
- Using *in vitro* methods developed for pharmaceuticals, we can make useful predictions of TK for large numbers of chemicals
- Exposure predictions and data are key to risk-based prioritization
  - Consensus modeling provides one path forward, but only as good as available data (at best)
  - New analytical chemistry tools (i.e., non-targeted analysis or NTA) may provide the data needed to understand what and how we are exposed to
- Exposure-based priority setting allows identification of the most relevant mixtures

United States Environmental Protection Agency

ExpoCast Project (Exposure Forecasting)

NCCT
Chris Grulke
Greg Honda*
Richard Judson
Ann Richard
Risa Sayre*
Mark Sfeir*
Rusty Thomas
John Wambaugh
Antony Williams

NRMRL
Xiaoyu Liu

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

NERL
Cody Addington*
Namdi Brandon*
Alex Chao*
Kathie Dionisio
Peter Egeghy
Hongtai Huang*
Kristin Isaacs
Ashley Jackson*
Jen Korol-Bexell*
Anna Kreutz*
Charles Lowe*
Seth Newton
Katherine Phillips
Paul Price
Jeanette Reyes*
Randolph Singh*
Marci Smeltz
Jon Sobus
John Streicher*
Mark Strynar
Mike Tornero-Velez
Elin Ulrich
Dan Vallero
Barbara Wetmore

***Trainees**

Collaborators

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

References

- Borassi, Michele, et al. "On the solvability of the six degrees of kevin bacon game." International Conference on Fun with Algorithms. Springer, Cham, 2014.
- Borgelt, Christian. "Frequent item set mining." Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 2.6 (2012): 437-456.
- Breyer, Stephen. Breaking the vicious circle: Toward effective risk regulation. Harvard University Press, 2009
- Collins, James J., and Carson C. Chow. "It's a small world." Nature 393.6684 (1998): 409.
- Diamond, Jared M. "Assembly of species communities." Ecology and evolution of communities (1975): 342-444.
- 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.
- Hertzberg, Robert P., and Andrew J. Pope. "High-throughput screening: new technology for the 21st century." Current opinion in chemical biology 4.4 (2000): 445-451.
- Hopkins, Brian. "Kevin Bacon and graph theory." Problems, Resources, and Issues in Mathematics Undergraduate Studies 14.1 (2004): 5-11.
- Judson, Richard, et al. "The toxicity data landscape for environmental chemicals." Environmental health perspectives 117.5 (2008): 685-695.
- Kaewkhaw, Rossukon, et al. "Treatment paradigms for retinal and macular diseases using 3-D retina cultures derived from human reporter pluripotent stem cell lines." Investigative ophthalmology & visual science 57.5 (2016): ORSFl1-ORSFl11.
- Kapraun, Dustin et al., "A Method for Identifying Prevalent Chemical Combinations in the US Population," Environmental Health Perspectives, 2017
- Kavlock, Robert, et al. "Update on EPA's ToxCast program: providing high throughput decision support tools for chemical risk management." Chemical research in toxicology 25.7 (2012): 1287-1302.
- 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.
- 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)
- Schmidt, Charles W. "TOX 21: new dimensions of toxicity testing." (2009): A348-A353.
- Tornero-Velez, Rogelio, Peter P. Egeghy, and Elaine A. Cohen Hubal. "Biogeographical analysis of chemical co-occurrence data to identify priorities for mixtures research." Risk Analysis: An International Journal 32.2 (2012): 224-236.
- Travers, Jeffrey, and Stanley Milgram. "An experimental study of the small world problem." Social Networks. Academic Press, 1977. 179-197.
- USGAO. "Toxic substances: EPA has increased efforts to assess and control chemicals but could strengthen its approach." (2013).
- Watts, Duncan J. Six degrees: The science of a connected age. WW Norton & Company, 2004.
- Watts, Duncan J., and Steven H. Strogatz. "Collective dynamics of 'small-world' networks." nature 393.6684 (1998): 440.
- 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.