EPA United States Environmental Protection Agency

# Using Computational Toxicology to Enable Risk-Based Chemical Safety Decision Making

*Richard Judson*
*U.S. EPA, National Center for Computational Toxicology*
*Office of Research and Development*

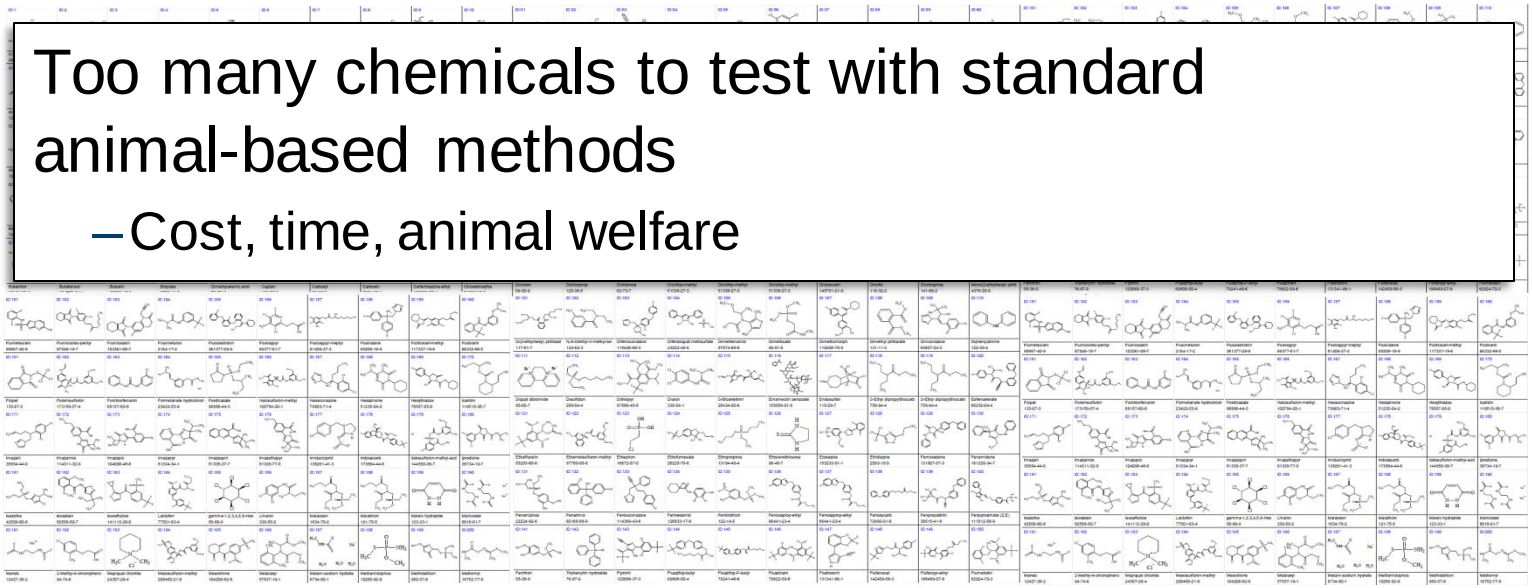

CSS Communities of Practice

November 17 2016

# Problem Statement

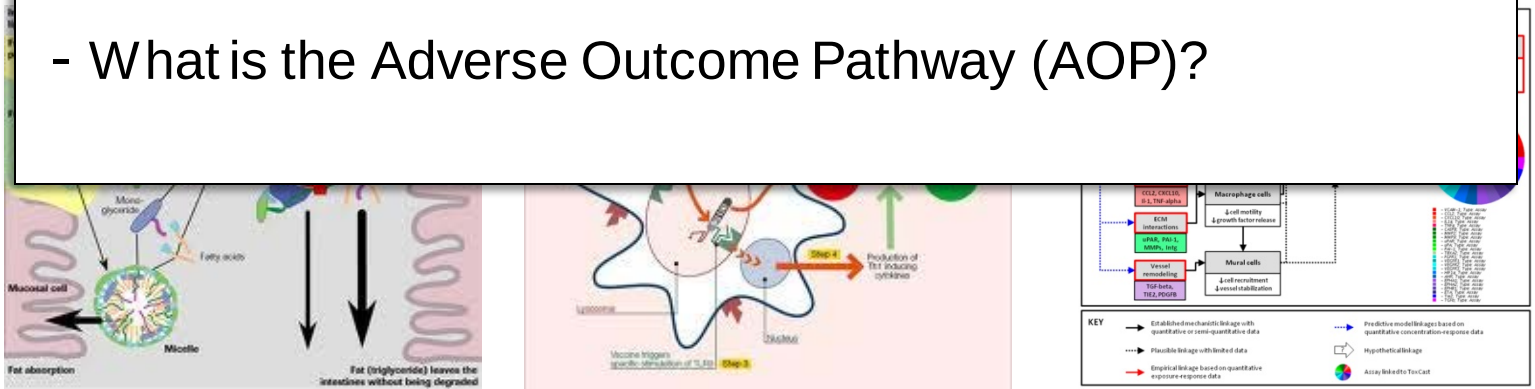
Too many chemicals to test with standard animal-based methods

- Cost, time, animal welfare



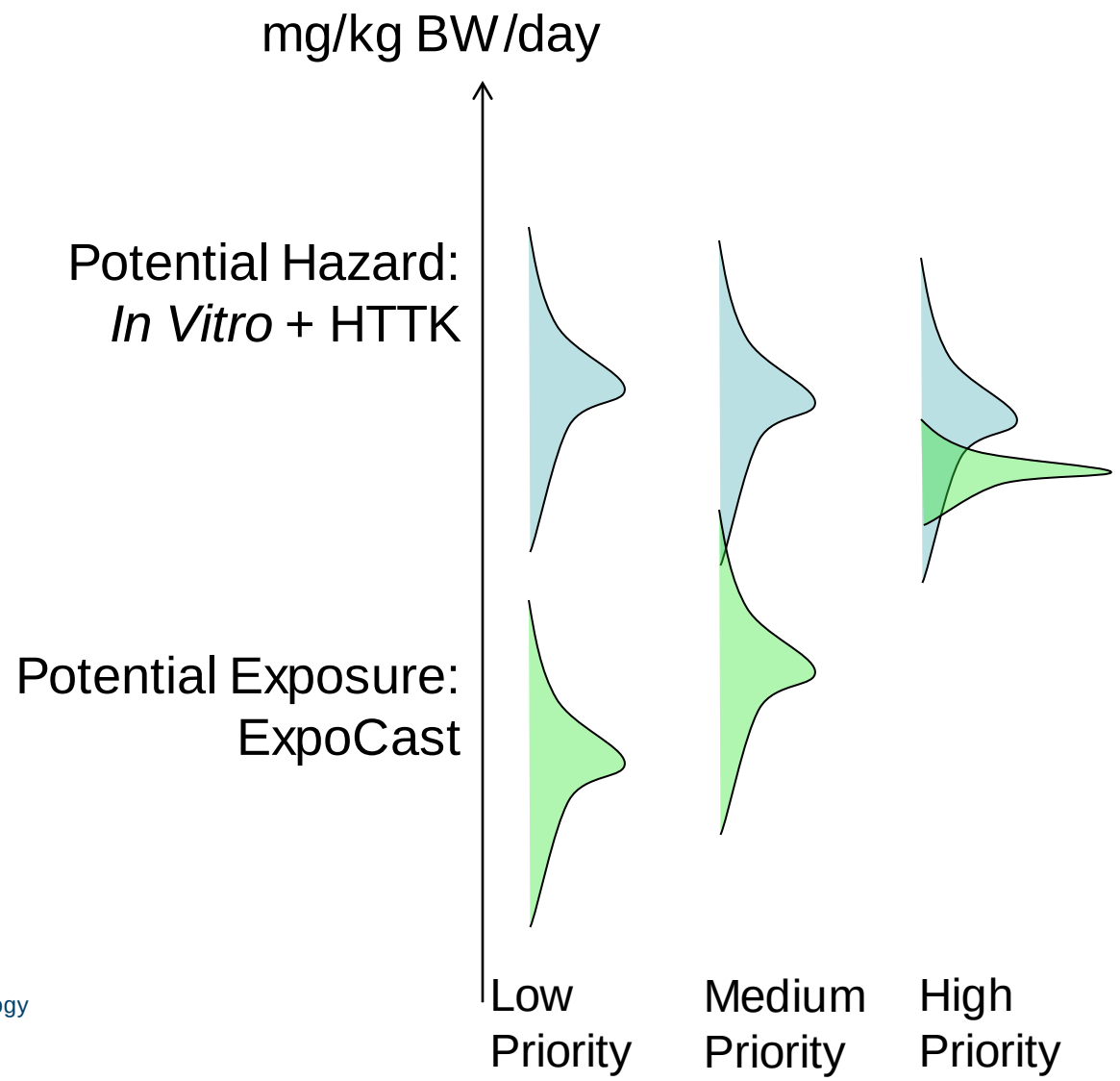
Need for better mechanistic data

- Determine human relevance
- What is the Adverse Outcome Pathway (AOP)?



# Risk-based Prioritization
# Hazard + Exposure

Semi-quantitative
*In Vitro* to *In Vivo*
Approach

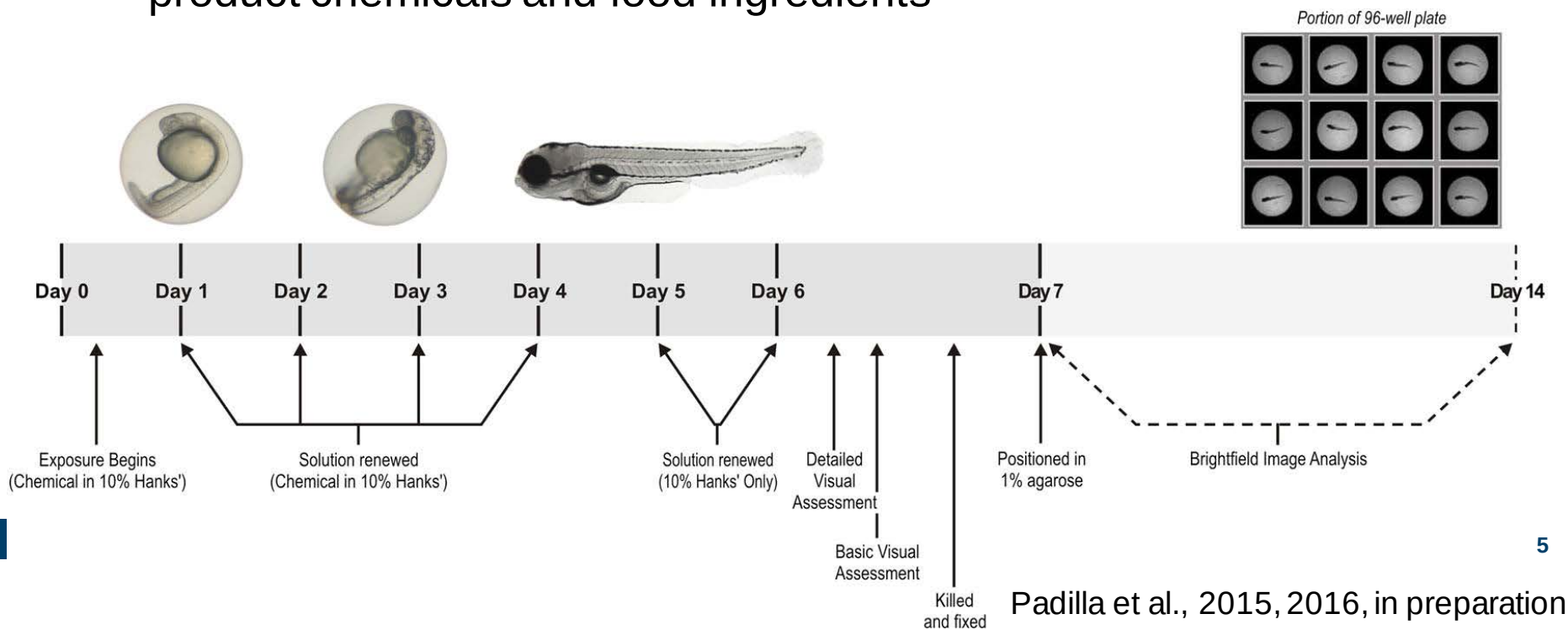


# Computational Toxicology

- Identify biological pathways of toxicity (AOPs)
- Develop high-throughput *in vitro* assays
  - Test "Human Exposure Universe" chemicals in the assays
- Develop models that link *in vitro* to *in vivo* hazard
  - Use pharmacokinetic models to predict activating doses
- Develop exposure models
- Add uncertainty estimates
- Create high-throughput risk assessments

# Zebrafish and Developmental Toxicology

- Goal: Use zebrafish as an *in vivo* model of vertebrate developmental toxicity
- Build *in vitro* to *in vivo* models using ~700 human assays
- ~1000 Chemicals
  - pharmaceuticals, pesticides, industrial chemicals, personal care product chemicals and food ingredients

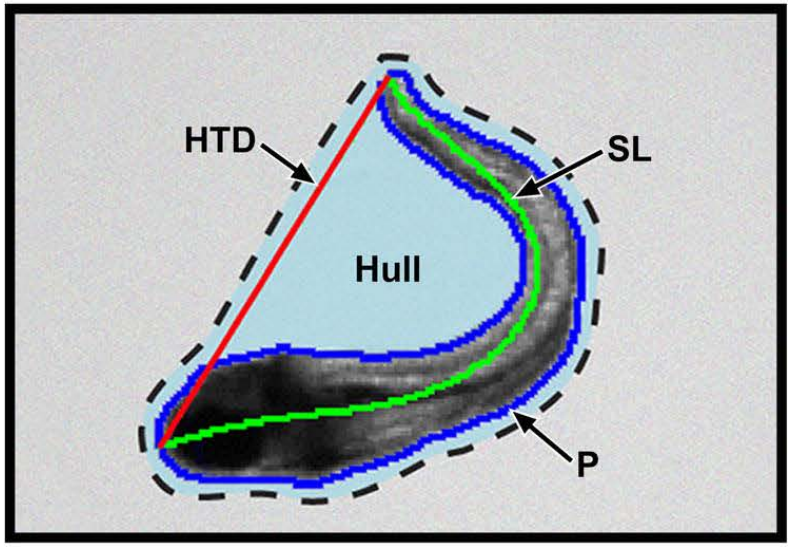





Padilla et al., 2015, 2016, in preparation

# Zebrafish Imaging and scoring

A

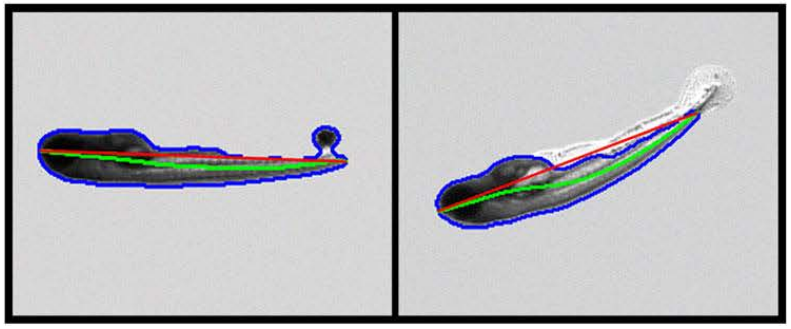



B

| Parameter | Description |
|---|---|
| Area | Area within the mask drawn around the fish, calculated as pixel count or micrometers |
| Perimeter–area (P) | A ratio of the outer perimeter of the fish to the area |
| SL | A line drawn approximately down the middle of the fish from the tip of the larvae's head to the tip of its tail |
| Width | The maximum distance perpendicular to the Spine Length |
| Length–width ratio | A ratio of SL to width |
| HTD | A direct line drawn from the tip of the larvae's head to the tip of the tail |
| Straightness | A ratio of HTD to SL |
| Convexity | A ratio of the fish area to the area of the hull |

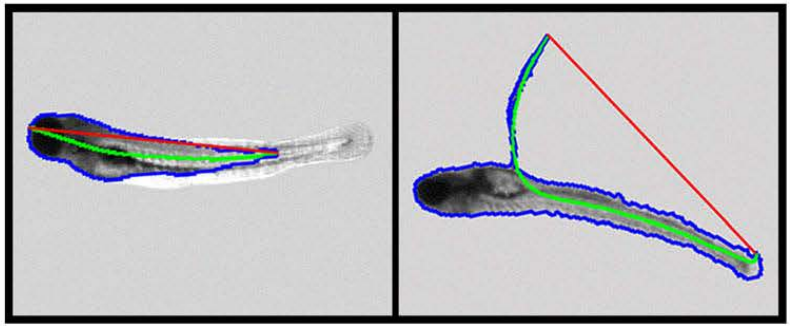
C

Acceptable



D

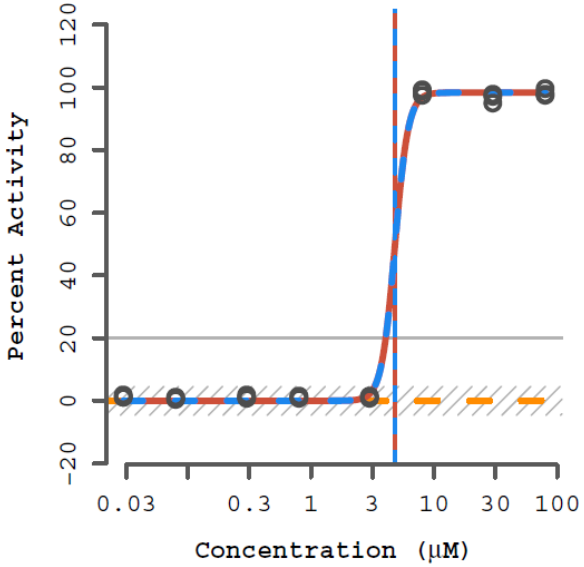
Unacceptable



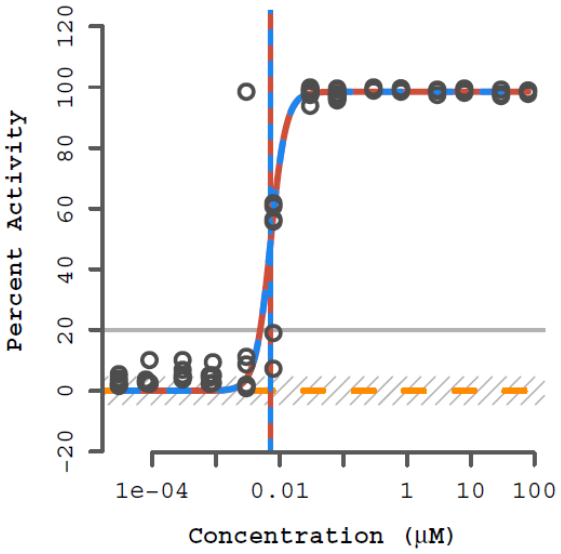
Deal et al. J Applied Tox. 2016

# Example chemicals

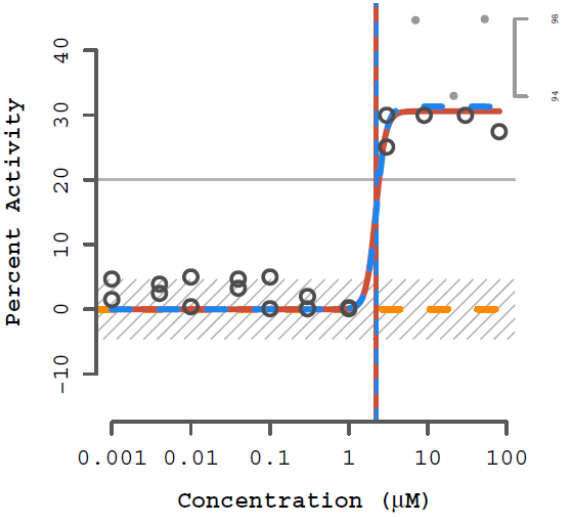
DES




Lovastatin

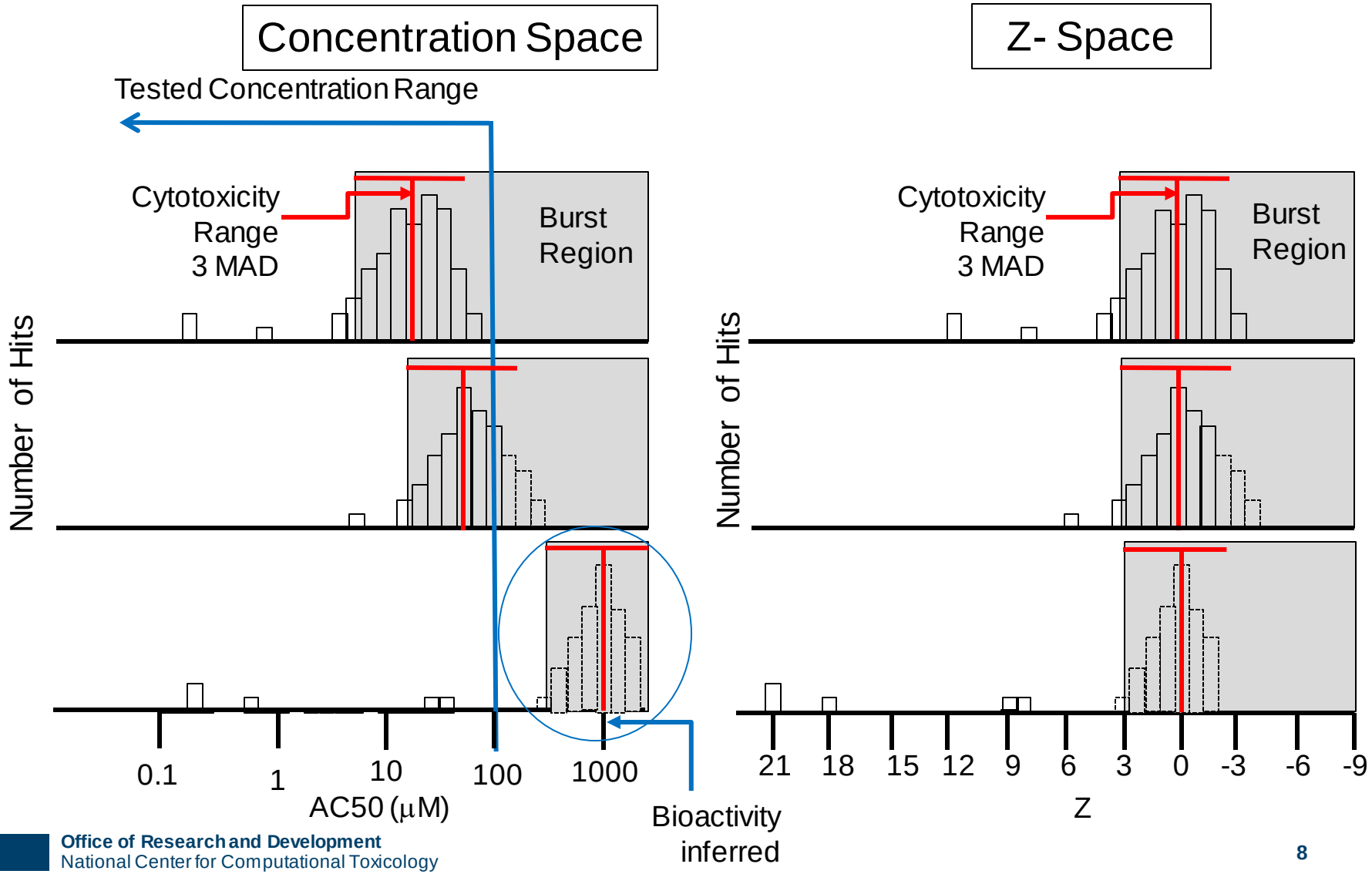



Permethrin




100% = death
<100% = malformations

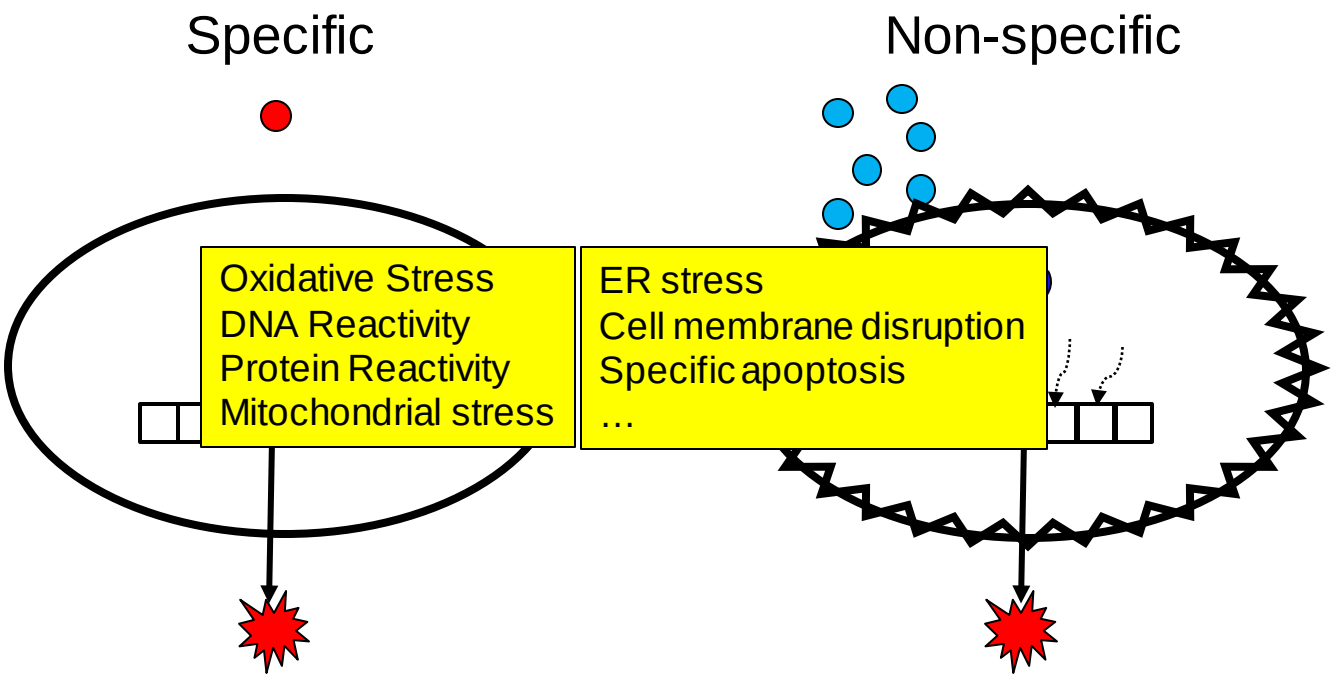
# Most chemicals display a "burst" of potentially non-selective bioactivity near cell-stress / cytotoxity conc.



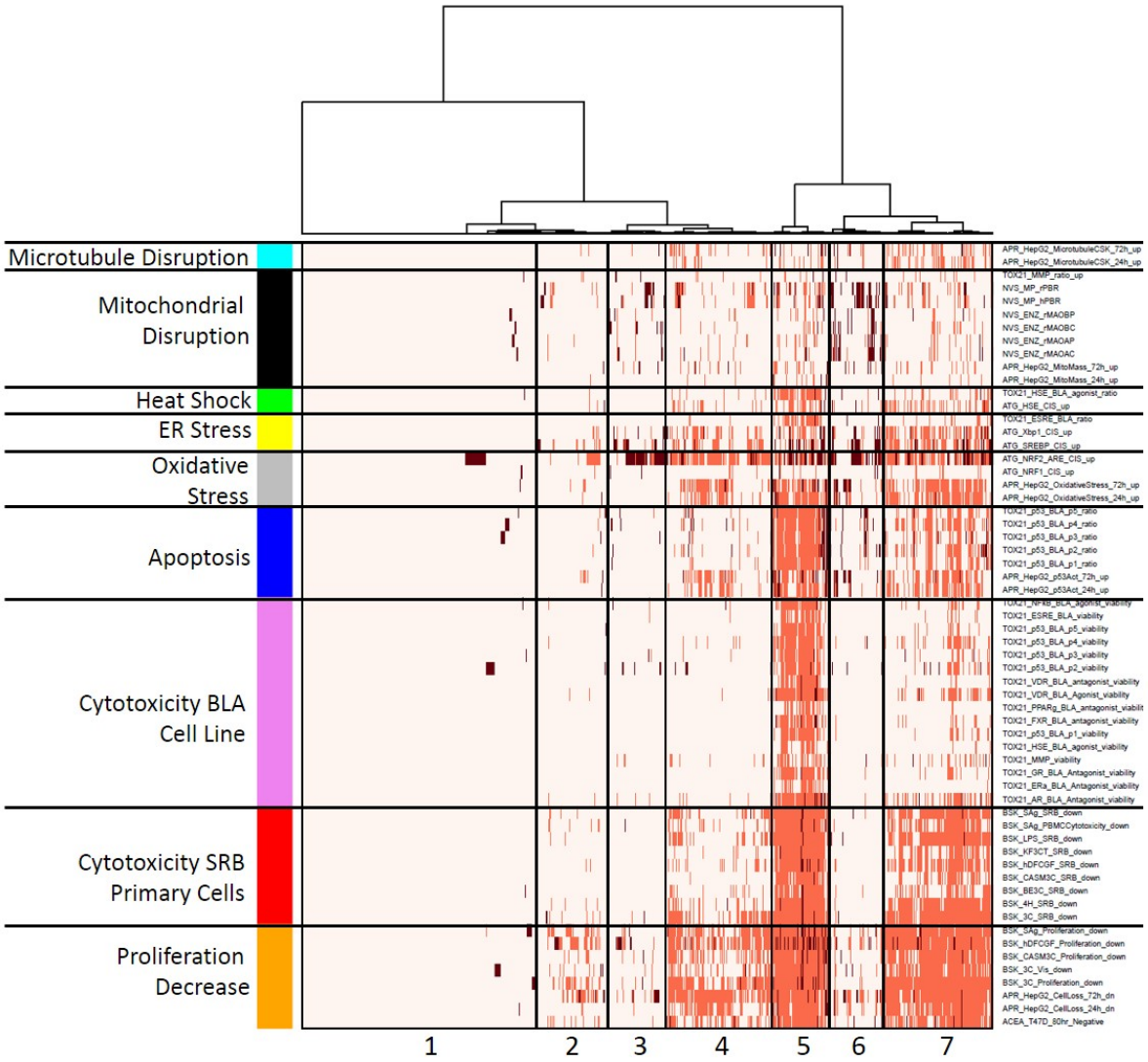



Judson et al. Tox.Sci. (2016)

# Schematic explanation of the burst



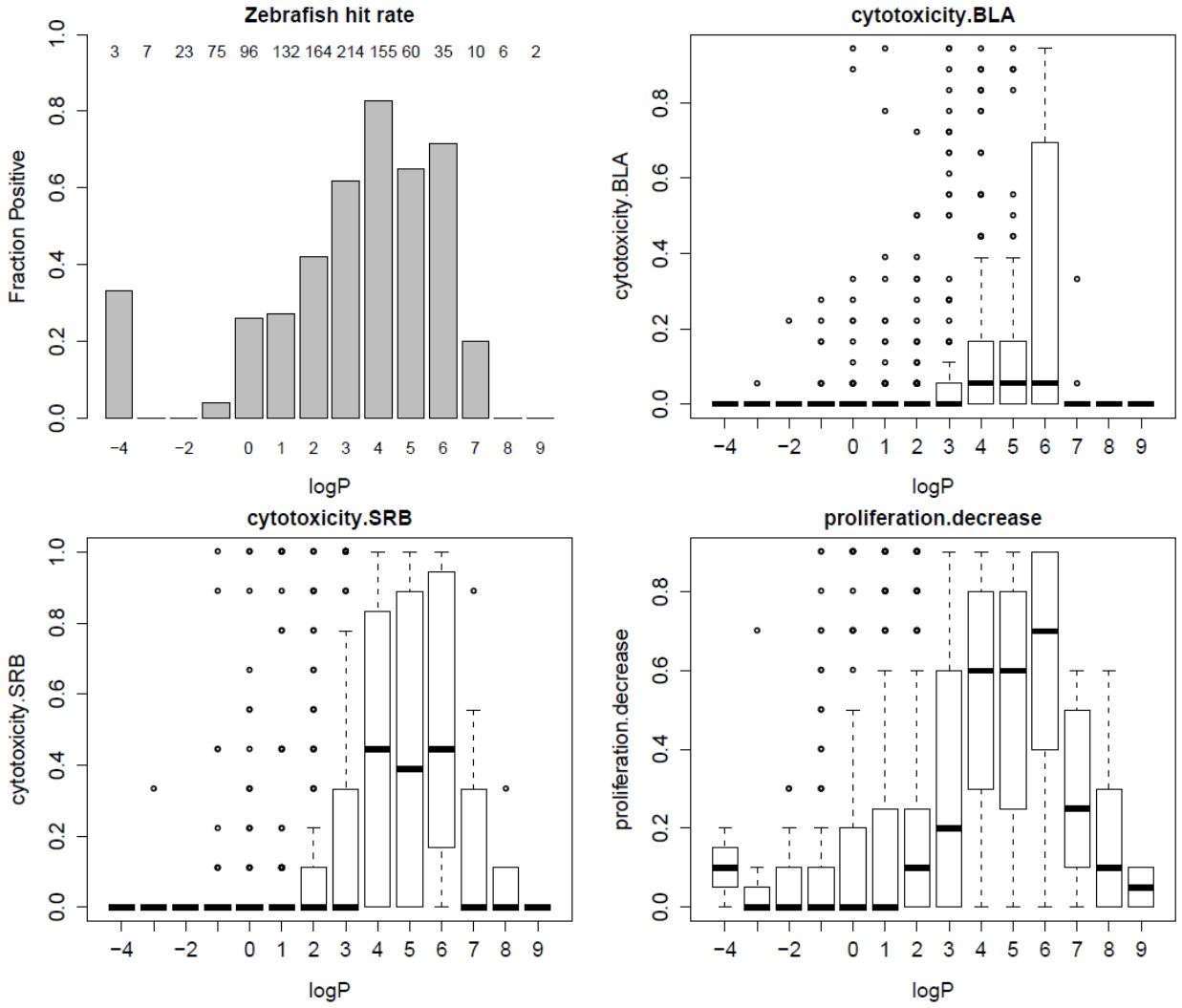
# Heatmap of stress and cytotoxicity assays in 1000 chemicals



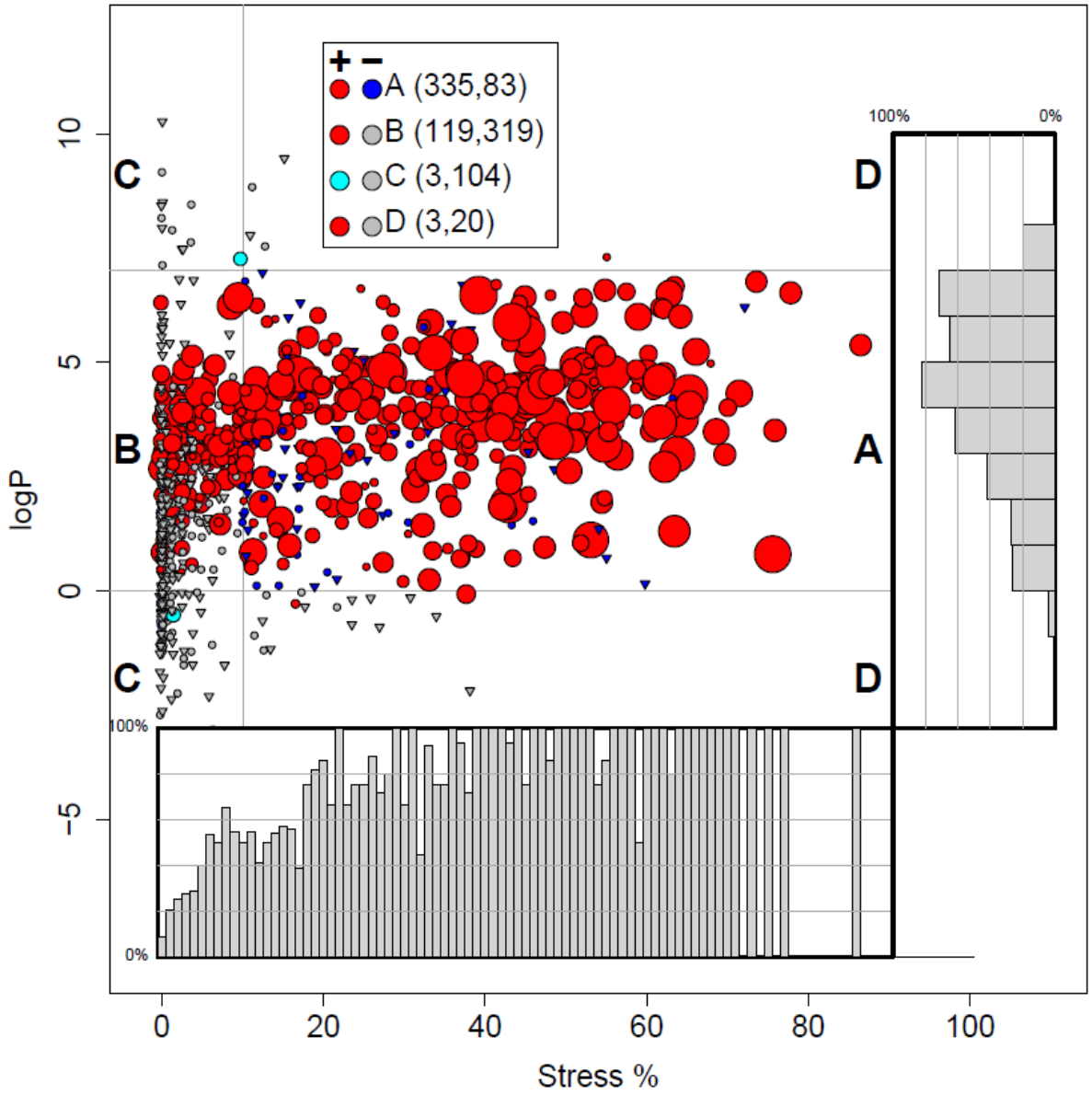

Judson et al. ToxSci (2016)

# Observation about logP

## Human *in vitro* cell stress behaves ~ zebrafish toxicity



# Stress, logP explains ~80% of ZF activity




ZF positive in conc-response

ZF negative in conc-response

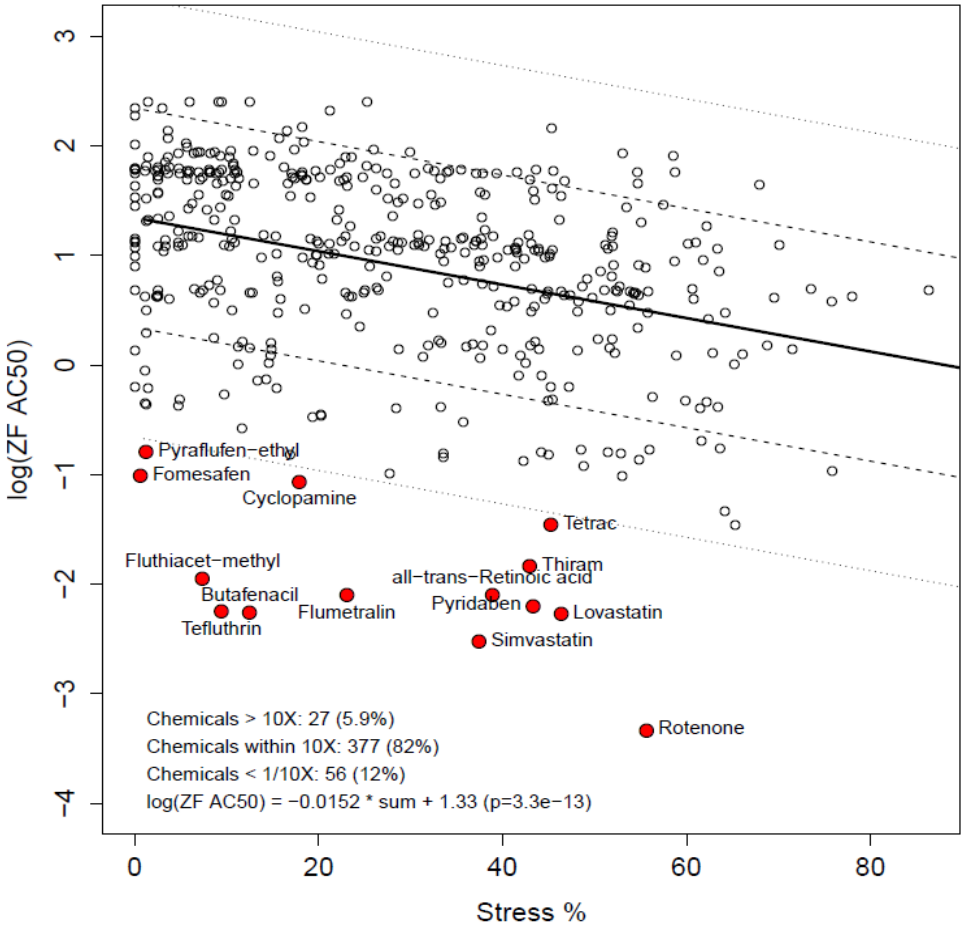
ZF negative in single conc

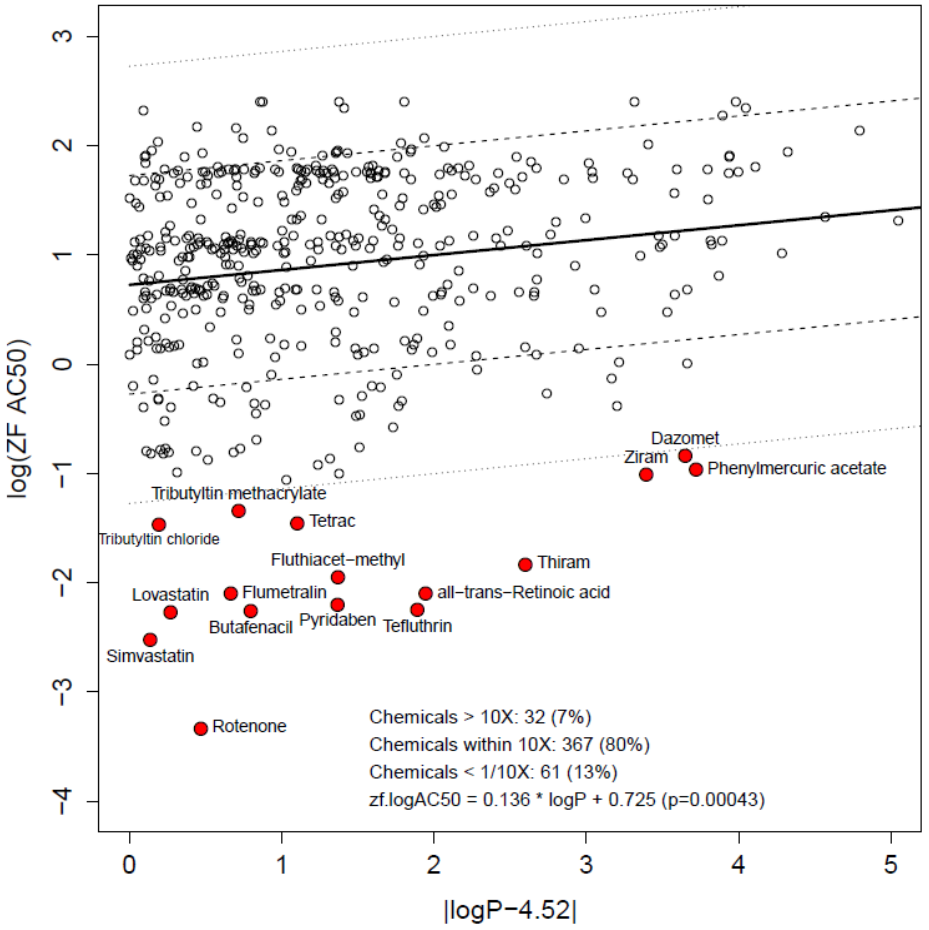
- 83 negatives in region A
  - Blue triangles
  - “false positives”?
- 50 “failed” single screen test?



Judson et al. In preparation

# “Excess Toxicity” points to specific target activity



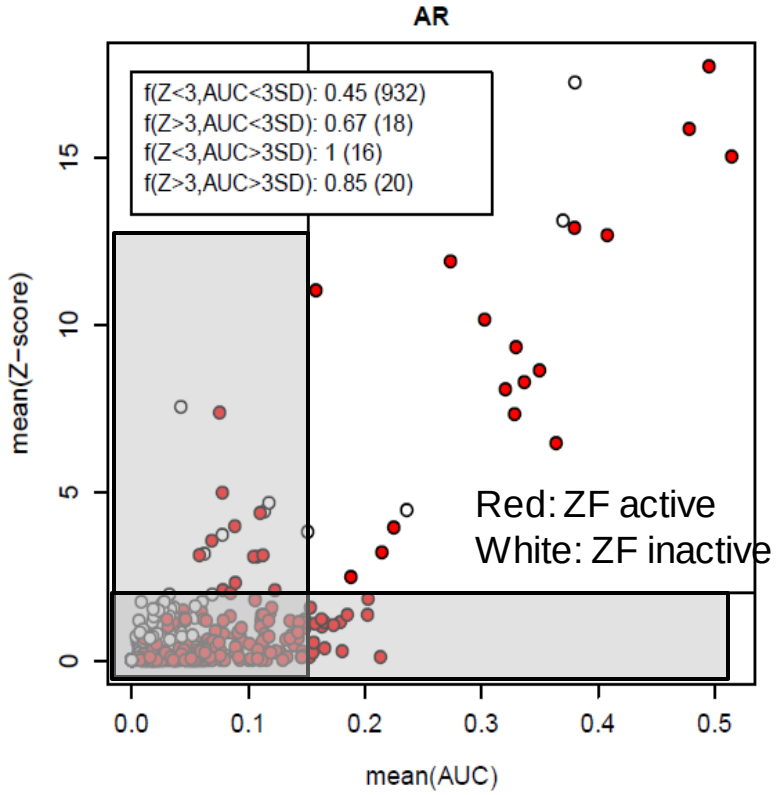
# Chemicals with excess toxicity tend to fall in a few target MOA classes

- ACHE
- Ion channel blockers
- HMGCR
- Mitochondrial disruptors
- PPO inhibitors (disrupts plant cell membranes)
- Chemicals reacting with protein SH groups
- Thyroid hormone receptor blockers

- Some of these classes are over-represented in overall hit predictivity and in excess potency for hits

# Look for specific targets by controlling for stress-related assay confounding

- Are potent actives against specific targets more likely than chance to be ZF-active?




Filter on Z-score (AC50 relative to cytotoxicity)

Filter on AUC (potency x efficacy)
Measure of reproducibility across multiple assays

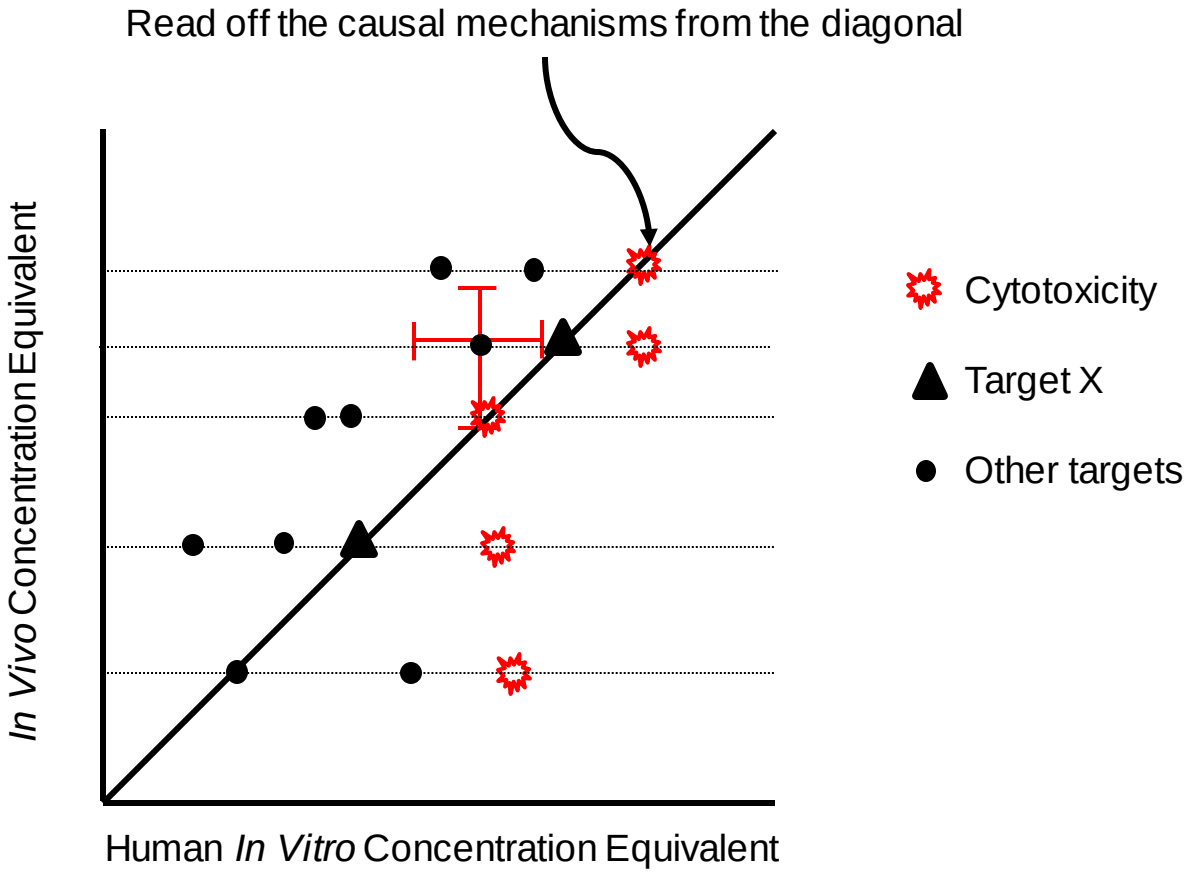
| class | Gene group | annotation | assays | TP | FP | FN | TN | Sens | Spec | BA | OR | PPV | p-value |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| endocrine | AR | Androgen receptor | 11 | 17 | 3 | 443 | 523 | 0.04 | 0.99 | 0.52 | 6.7 | 0.85 | 0.0005 |
| endocrine | CYP19A1 | Aromatase | 2 | | | | | | | 0.52 | 14.4 | 0.92 | 9E-07 |
| endocrine | ESR | Estrogen receptor | 1 | | | | | | | 0.53 | 5.8 | 0.83 | 2E-05 |
| endocrine | NR3C1 | Glucocorticoid receptor | 4 | 14 | 4 | 446 | 522 | 0.03 | 0.99 | 0.51 | 4.1 | 0.78 | 0.0084 |
| endocrine | PGR | Progesterone receptor | 2 | 15 | 3 | 445 | 523 | 0.03 | 0.99 | 0.51 | 5.9 | 0.83 | 0.0016 |
| ER stress | SREBF1 | | 1 | 36 | 10 | 424 | 516 | 0.08 | 0.98 | 0.53 | 4.4 | 0.78 | 1E-05 |
| ER stress | XBP1 | | 1 | 10 | 1 | 450 | 525 | 0.02 | 1.00 | 0.51 | 11.7 | 0.91 | 0.0039 |
| GPCR | LTD4 | | 1 | 11 | 1 | 449 | 525 | 0.02 | 1.00 | 0.51 | 12.9 | 0.92 | 0.002 |
| growth factor | EGR1 | | 1 | 19 | 1 | 441 | 525 | 0.04 | 1.00 | 0.52 | 22.6 | 0.95 | 8E-06 |
| hypoxia | HIF1A | | 1 | 24 | 3 | 436 | 523 | 0.05 | 0.99 | 0.52 | 9.6 | 0.89 | 5E-06 |
| inflammation | CEBPB | | 1 | 30 | 6 | 430 | 520 | 0.07 | 0.99 | 0.53 | 6.0 | 0.83 | 5E-06 |
| inflammation | CREB3 | | 1 | 23 | 1 | 437 | 525 | 0.05 | 1.00 | 0.52 | 27.6 | 0.96 | 5E-07 |
| inflammation | PTGER2 | | 1 | | | | | | | | 5.0 | 0.81 | 3E-05 |
| inflammation | TNF | | 1 | | | | | | | | 2.8 | 0.70 | 0.0026 |
| ion channel | KCNH2 | | 1 | | | | | | | | 7.6 | 0.87 | 0.0026 |
| oncogene | JUN | | 1 | 18 | 6 | 442 | 520 | 0.04 | 0.99 | 0.51 | 3.5 | 0.75 | 0.0062 |
| oxidative stress | NFE2L2 | NRF2, ROS Sensor | 2 | 34 | 5 | 426 | 521 | 0.07 | 0.99 | 0.53 | 8.3 | 0.87 | 1E-07 |
| transcription factor | POU2F1 | | 1 | 17 | 4 | 443 | 522 | 0.04 | 0.99 | 0.51 | 5.0 | 0.81 | 0.0016 |
| transcription factor | SMAD1 | | 1 | 21 | 5 | 439 | 521 | 0.05 | 0.99 | 0.52 | 5.0 | 0.81 | 0.0005 |
| transcription factor | SOX1 | | 1 | 16 | 5 | 444 | 521 | 0.03 | 0.99 | 0.51 | 3.8 | 0.76 | 0.0072 |
| transcription factor | SP1 | | 1 | 18 | 2 | 442 | 524 | 0.04 | 1.00 | 0.52 | 10.7 | 0.90 | 6E-05 |
| transporter | DAT | | 1 | 18 | 6 | 442 | 520 | 0.04 | 0.99 | 0.51 | 3.5 | 0.75 | 0.0062 |
| xenobiotic metabolism | CYP1A | cytochrome P450 | 4 | 18 | 3 | 442 | 523 | 0.04 | 0.99 | 0.52 | 7.1 | 0.86 | 0.0003 |
| xenobiotic metabolism | CYP2A | cytochrome P450 | 3 | 25 | 5 | 435 | 521 | 0.05 | 0.99 | 0.52 | 6.0 | 0.83 | 5E-05 |
| xenobiotic metabolism | CYP2B | cytochrome P450 | 2 | 25 | 2 | 435 | 524 | 0.05 | 1.00 | 0.53 | 15.1 | 0.93 | 4E-07 |
| xenobiotic metabolism | CYP2C | cytochrome P450 | 8 | | | | | | | | 1E+06 | 1.00 | 8E-09 |
| xenobiotic metabolism | CYP2D | cytochrome P450 | 3 | | | | | | | | 5.9 | 0.83 | 0.0016 |
| xenobiotic metabolism | CYP2J | cytochrome P450 | 1 | 21 | 1 | 439 | 525 | 0.05 | 1.00 | 0.52 | 25.1 | 0.95 | 2E-06 |
| xenobiotic metabolism | CYP3A | cytochrome P450 | 4 | 19 | 1 | 441 | 525 | 0.04 | 1.00 | 0.52 | 22.6 | 0.95 | 8E-06 |
| xenobiotic metabolism | NR1I2 | PXR | 3 | 30 | 9 | 430 | 517 | 0.07 | 0.98 | 0.52 | 4.0 | 0.77 | 0.0001 |

Endocrine pathways

Largely stress activity:
more potent than cytotoxicity

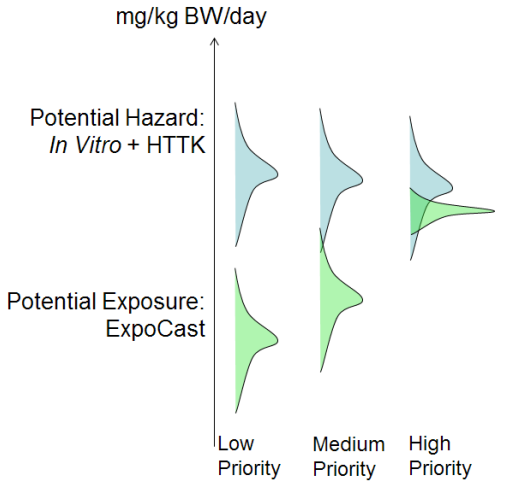
Largely due to conazoles

# The ideal *in vitro* to *in vivo* model
# Zebrafish, rat, mouse, human, …




- Failure so far – concentration equivalents require better understanding of relative kinetics, bioavailability
- Also concentration uncertainty on both axes is ~1 log unit (95% CI)

# Modeling with Uncertainty




- Our first goal is **prediction**
  - What is the highest safe dose of a chemical?
  - What types of harm would a chemical cause above that dose?
- Predictions are based on **models**
  - Computational, statistical, "mental", *in vitro*, *in vivo*
- All models are based on **data**
- Data is always subject to noise, variability
- Therefore, all predictions are subject to **uncertainty**

- Our second goal is estimating **prediction uncertainty**

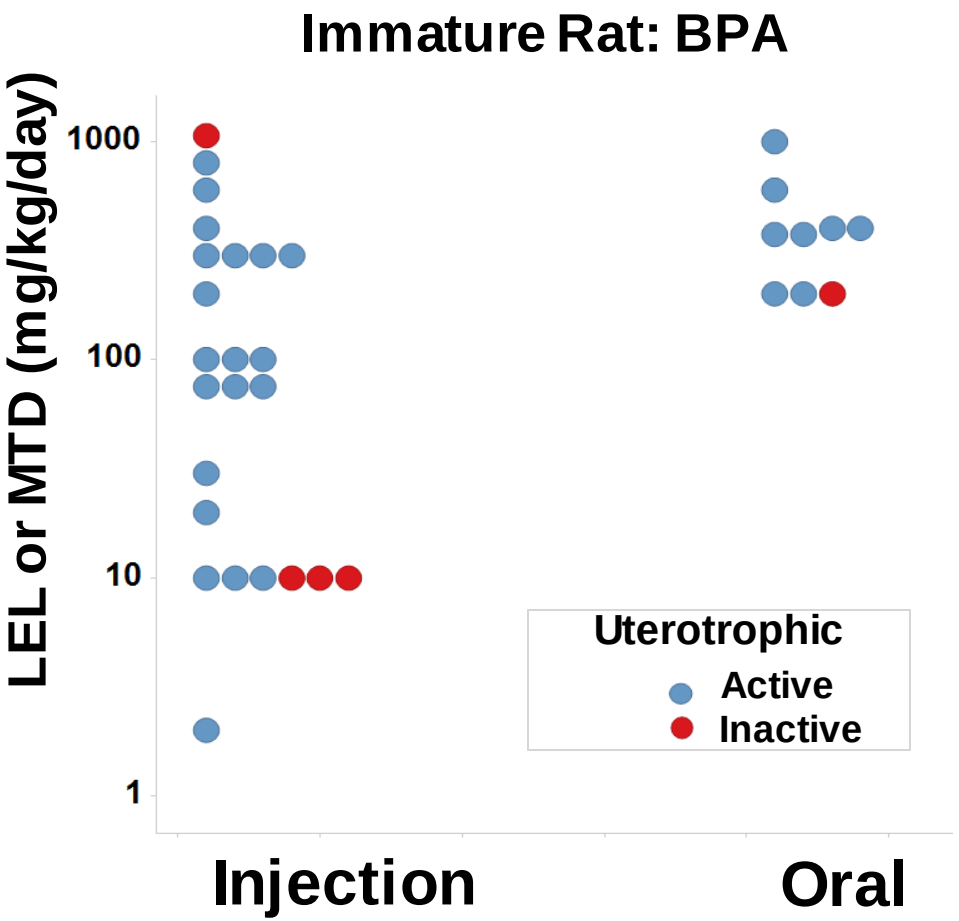


Watt, Kapraun et al. In preparation

# *In vivo* guideline study uncertainty

## 26% of chemicals tested multiple times in the uterotrophic assay gave discrepant results

United States Environmental Protection Agency




Kleinstreuer et al. EHP 2015

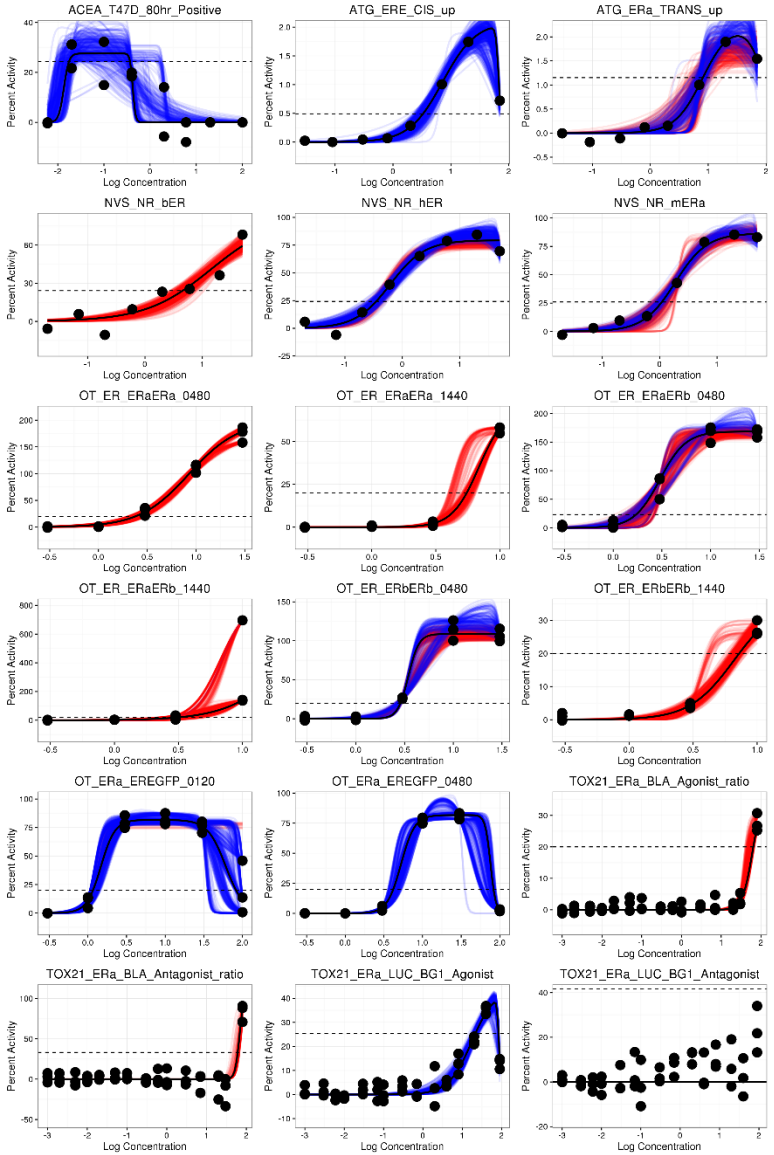
### Anemia Reproducibility

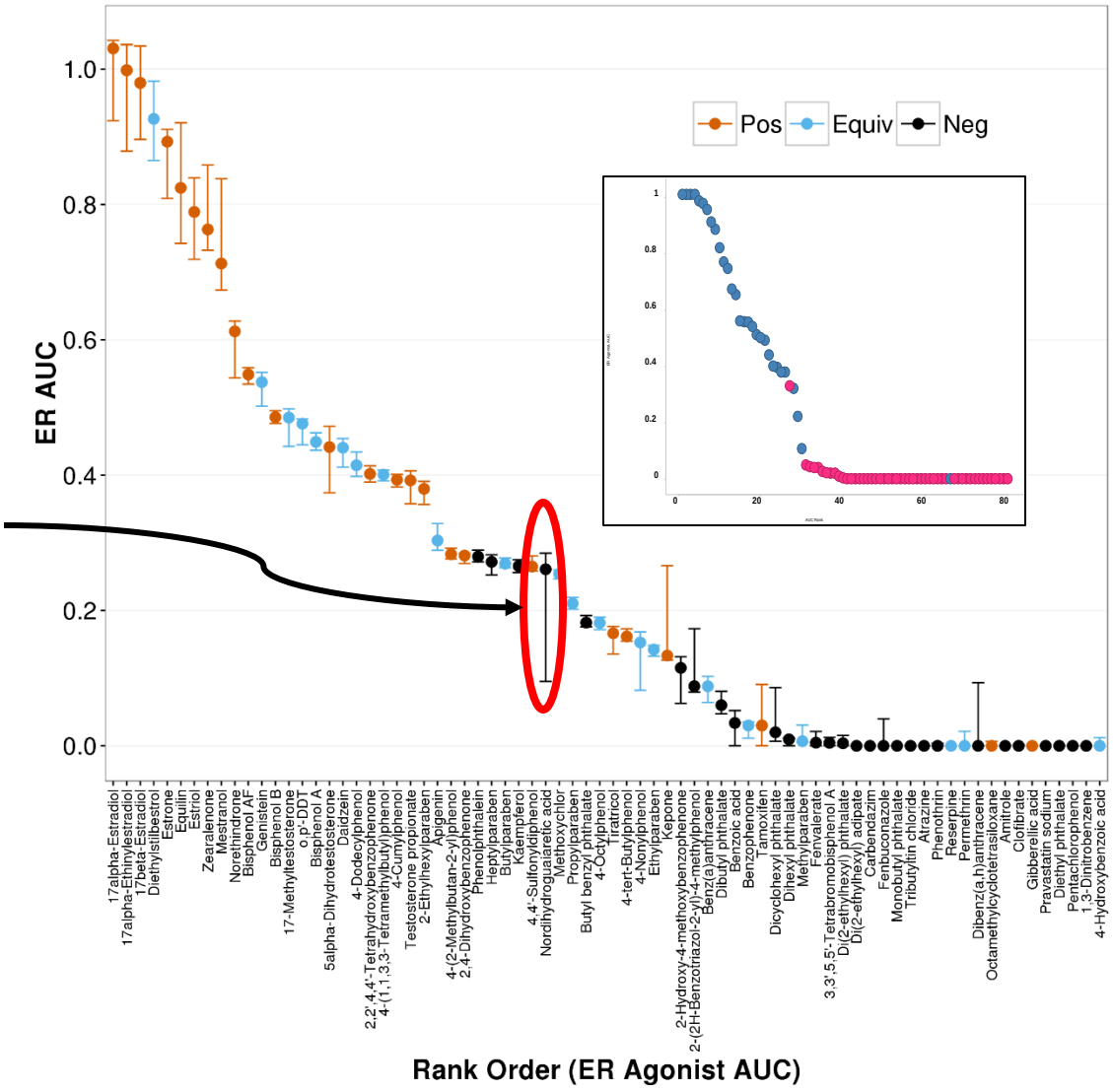
| species / study 1 | species / study 2 | Reproduce | Does Not Reproduce | Fraction Reproduce |
|---|---|---|---|---|
| rat SUB | rat CHR | 18 | 2 | 0.90 |
| rat CHR | dog CHR | 13 | 2 | 0.87 |
| rat CHR | rat SUB | 18 | 4 | 0.82 |
| rat SUB | rat SUB | 16 | 4 | 0.80 |
| rat SUB | dog CHR | 11 | 4 | 0.73 |
| mouse CHR | rat CHR | 11 | 4 | 0.73 |
| mouse CHR | rat SUB | 13 | 7 | 0.65 |
| dog CHR | rat SUB | 11 | 6 | 0.65 |
| dog CHR | rat CHR | 13 | 8 | 0.62 |
| rat CHR | mouse CHR | 11 | 11 | 0.50 |
| mouse CHR | dog CHR | 6 | 6 | 0.50 |
| rat SUB | mouse CHR | 13 | 14 | 0.48 |
| dog CHR | mouse CHR | 6 | 8 | 0.43 |
| mouse CHR | mouse CHR | 2 | 3 | 0.40 |

Judson et al. In Preparation

# *In Vitro* Assay Data is also subject to uncertainty

## See Eric Watt poster



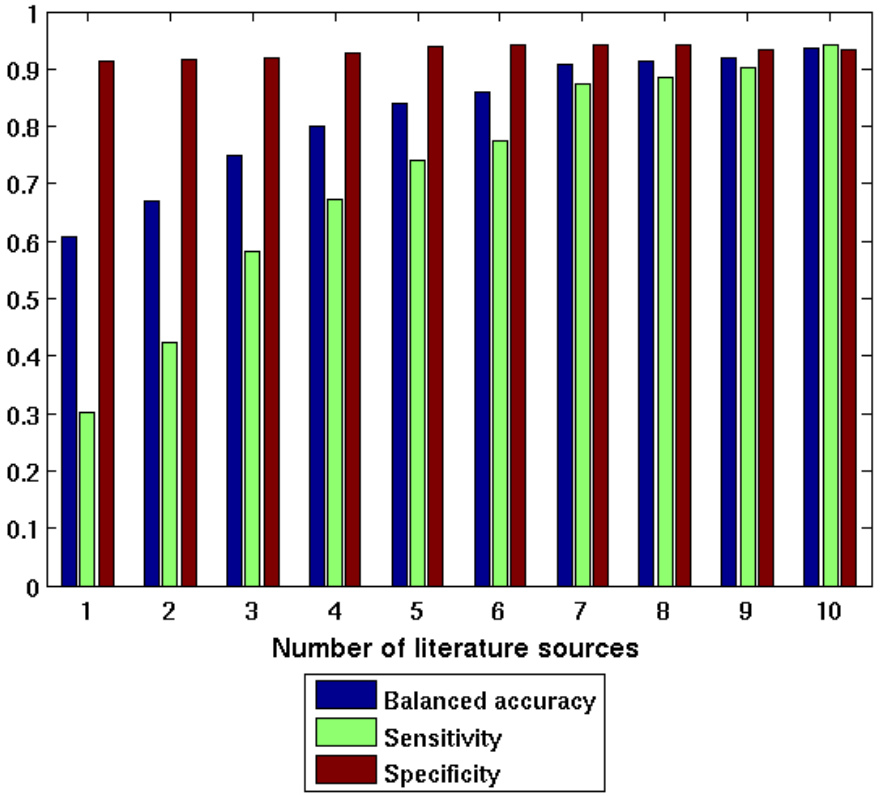



Watt et al. (in prep)

# Uncertainty in data has big impact on model performance

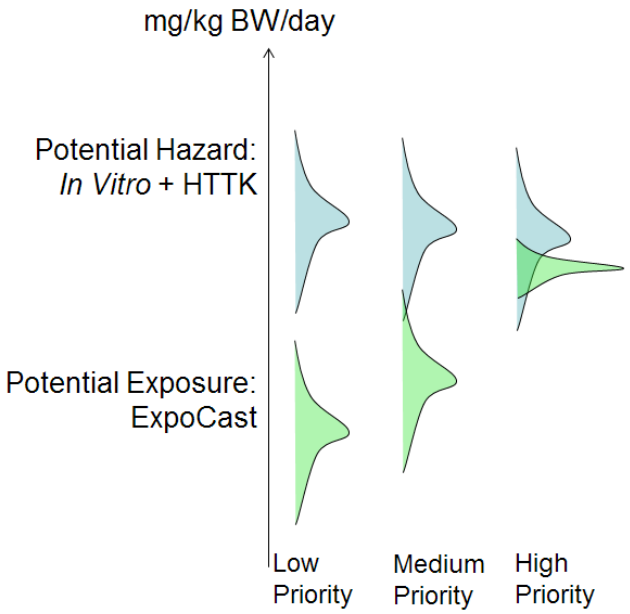
As greater consistency is required from literature sources, QSAR consensus model performance improves

- Source: CERAPP project, Mansouri et al. EHP 2015
- Community development of estrogen receptor models tested against thousands of experimental data points

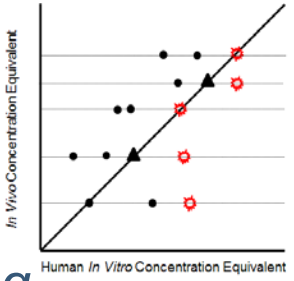


# Given all the uncertainty, is modeling futile?

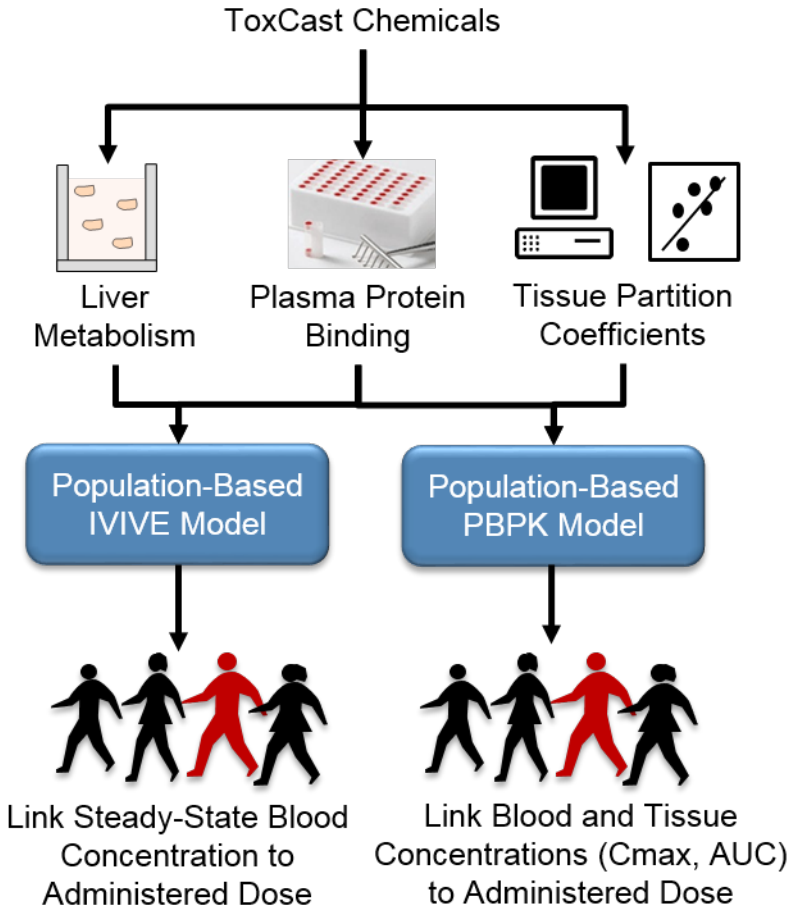
- Not in risk assessment
  - What's important is the difference between hazard and exposure
- Hazard Model:
  - In vitro IC50 (μM) with uncertainty
  - Use toxico / pharmacokinetic model to convert to mg/kg/day (with added uncertainty)
- Exposure model
  - Based on NHANES, other biomonitoring data
  - Add uncertainty
- Compare ranges for margin of exposure

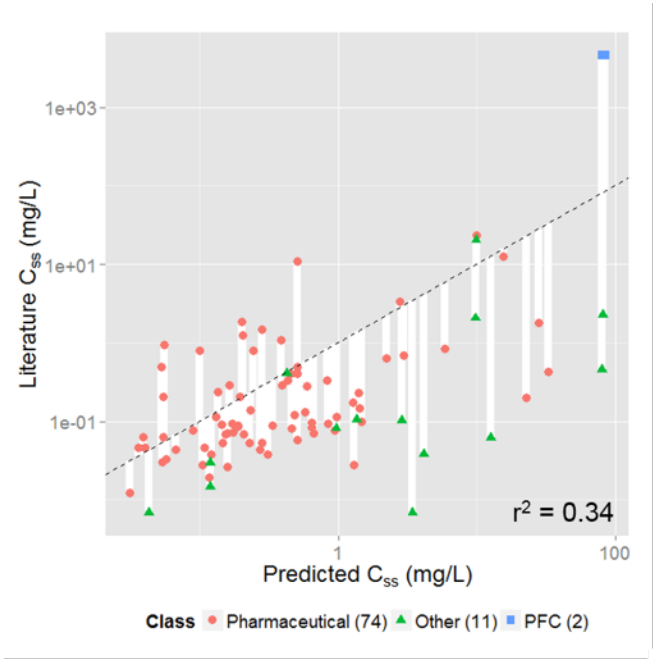


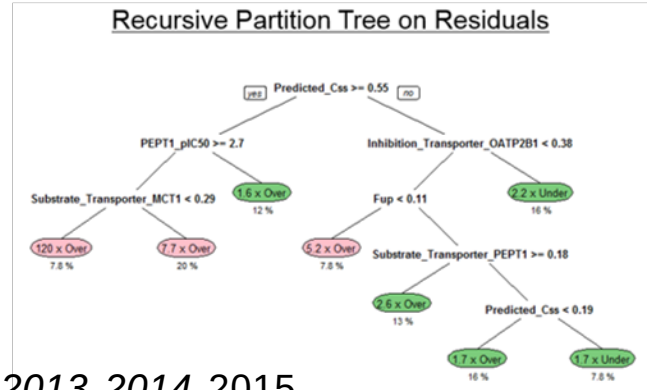
# Toxicokinetics Modeling




## *Incorporating Dosimetry and Uncertainty into In Vitro Screening*



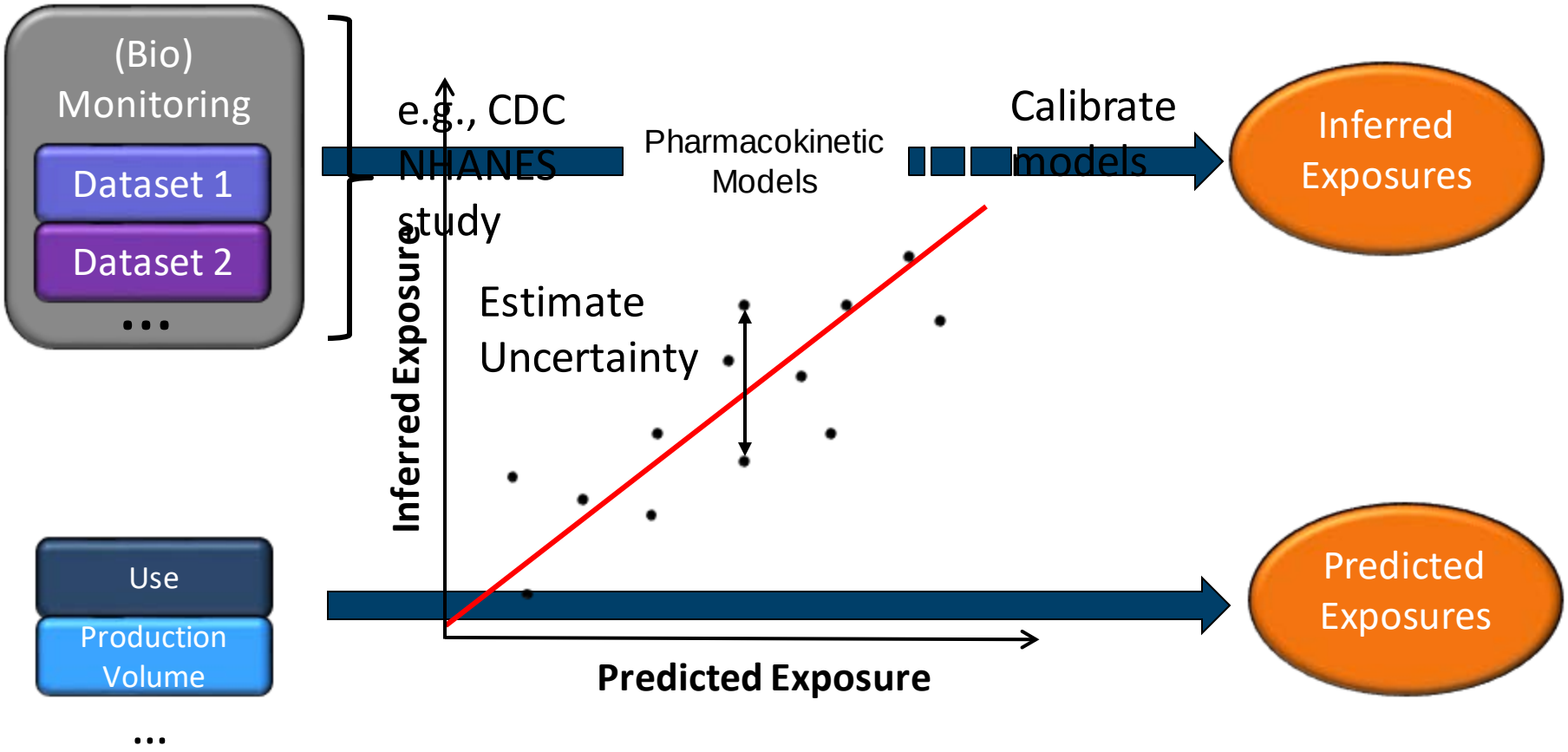






Wetmore, Rotroff, Wambaugh *et al., 2013, 2014,* 2015

# Population and Exposure Modeling

*Estimating Exposure and Associated Uncertainty with Limited Data*



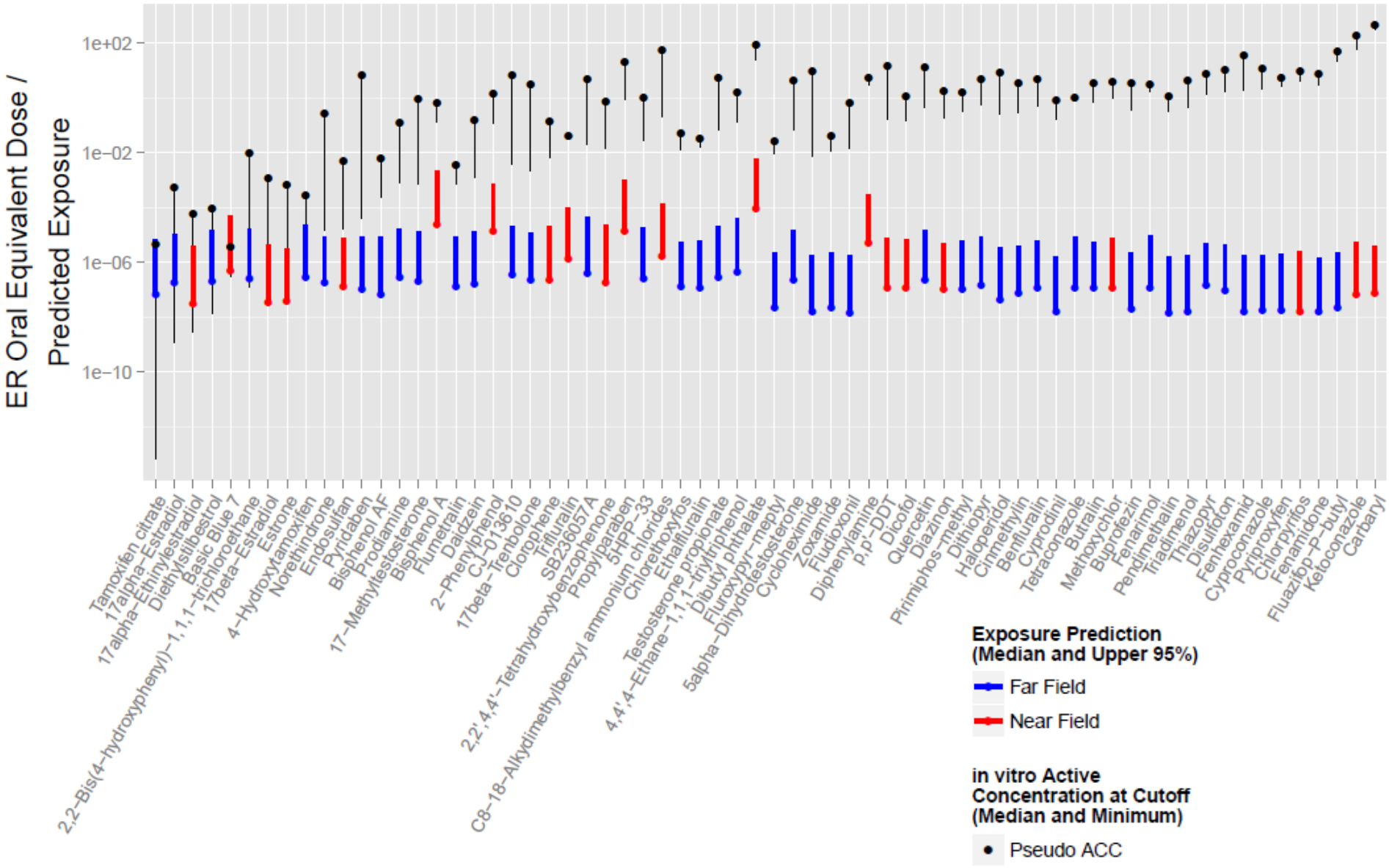





Wambaugh *et al.*, 2014

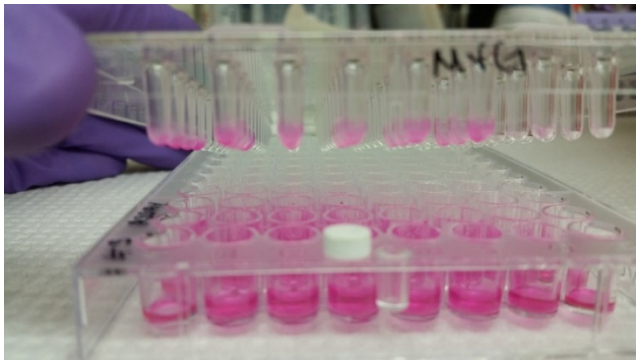
EPA United States Environmental Protection Agency

# High-throughput Risk Assessment for ER
# 290 chemicals with ER bioactivity

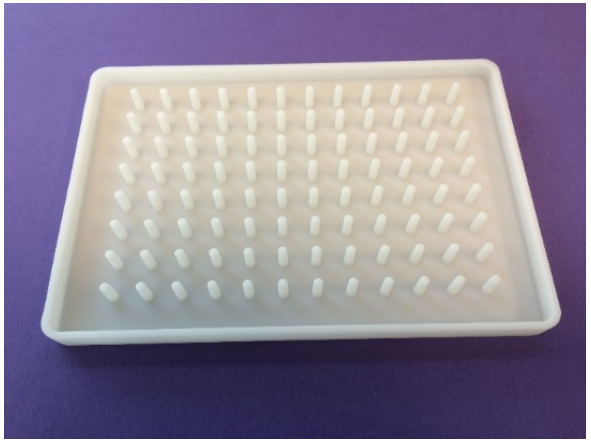


# Retrofitting Assays for Metabolic Competence – Extracellular Approach

Alginate Immobilization of Metabolic Enzymes (AIME)

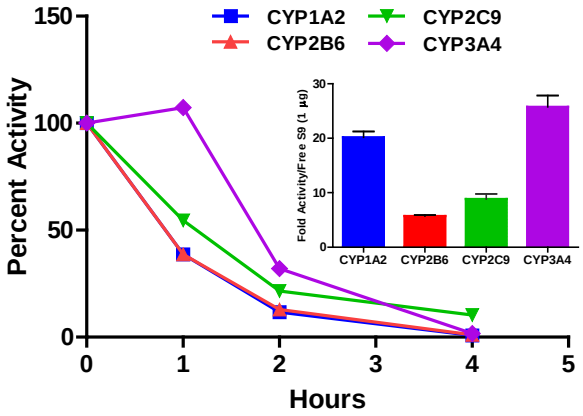


Prototype Lids

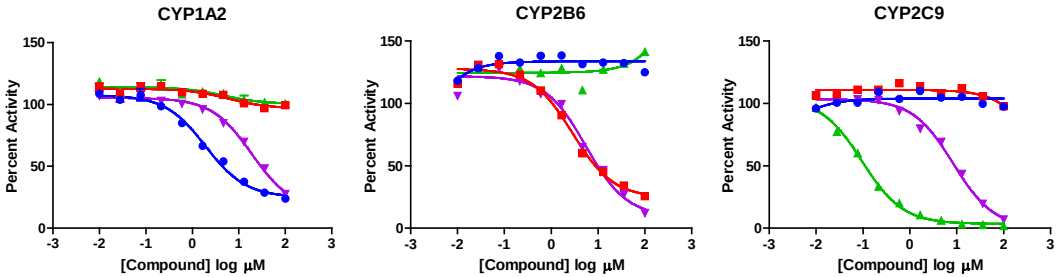


DeGroot et al. 2016 SOT poster #3757

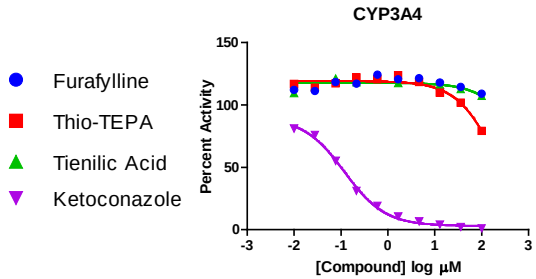
Amount of XME Activity in Microspheres



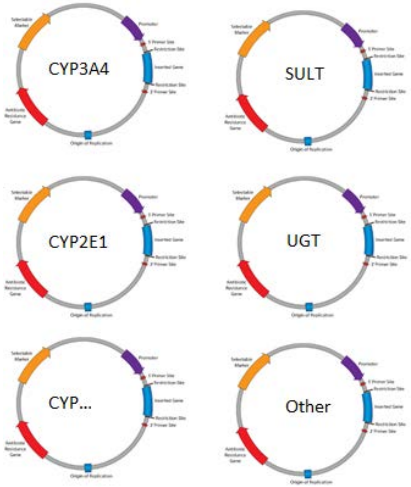

Small Molecule Inhibition of XME Activity



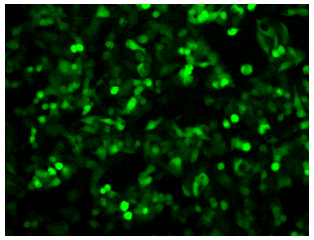

| Compound | Mol. Wt. (g/mol) | Targeted P450 | IC50 Free S9 (µM) | IC50 AIME (µM) |
|---|---|---|---|---|
| Furafylline | 260.25 | 1A2 | 2.39 | 1.92 |
| Thio-TEPA | 189.22 | 2B6 | 7.46 | 2.86 |
| Tienilic Acid | 331.17 | 2C9 | .053 | .096 |
| Ketoconazole | 531.43 | 3A4 | .086 | 0.12 |



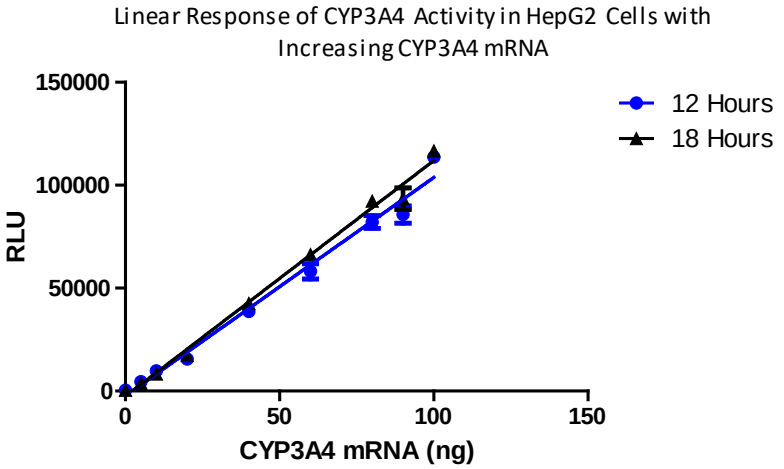
# Retrofitting Assays for Metabolic Competence – mRNA Intracellular Strategy




Pool in vitro transcribed mRNAs chemically modified with pseudouridine ad 5-methylcytidine to reduce immune stimulation

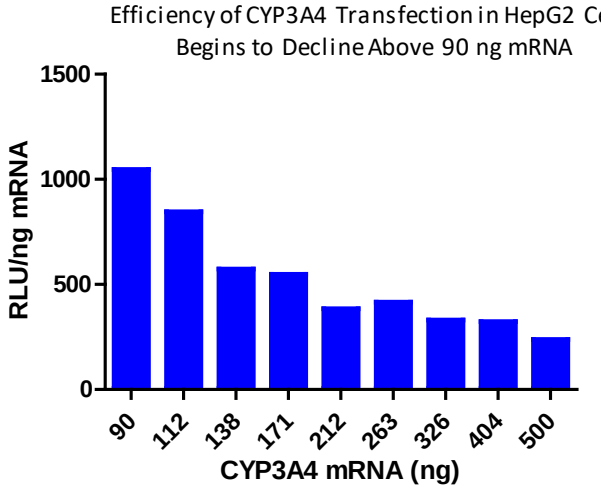


293T cells 21.5 h post transfection with 90 ng of EGFP mRNA using TransIT reagent




Advantage of transfecting with mRNA

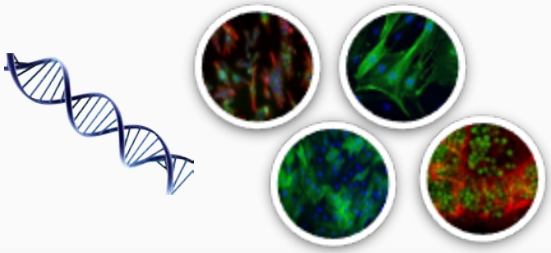
Titrate different CYPs to match different ratios in different tissues



# Developing Approaches for Tiered Testing

United States Environmental Protection Agency

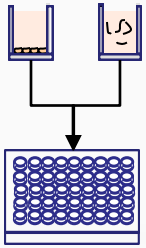
Comprehensive Transcriptomic Screening

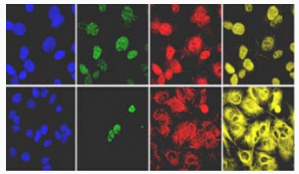


Multiple Human Cell Types

Comprehensive Characterization

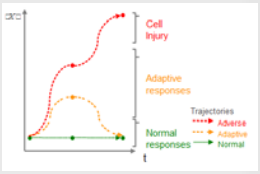
Focused ToxCast/Tox21 Assays

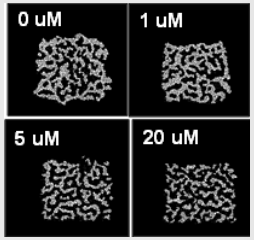




Time Course High Content Assays

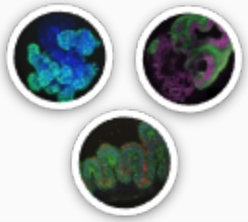
Verification of Affected Processes/ Pathways and Temporal Evaluation




Computational and Statistical Modeling

Organs-on-a-Chip



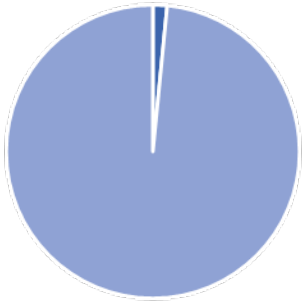
Organotypic and Organoid Models

Interpretation of Affected Process/ Pathways and Population Variability

# Planning for HT Transcriptomics

*New Approaches to Comprehensively Assess Potential Biological Effects*

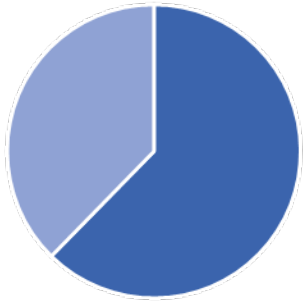
**Gene Coverage**



ToxCast

Not in ToxCast

**Pathway Coverage***

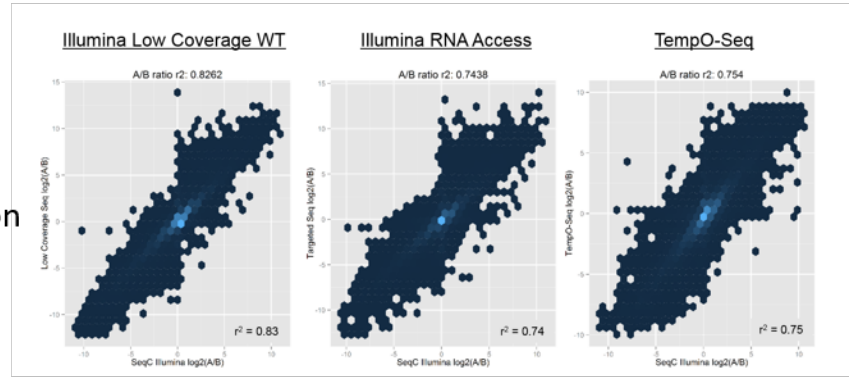


*At least one gene from pathway represented

High-Throughput Transcriptomic Platforms

- Low-cost
- Whole genome
- 384-well
- Automatable

Technical Comparison

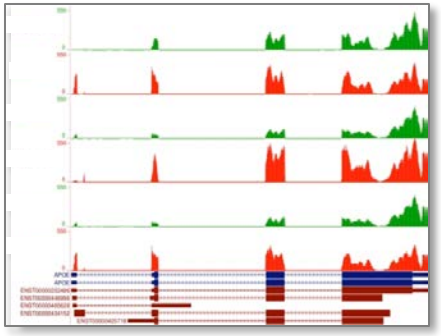


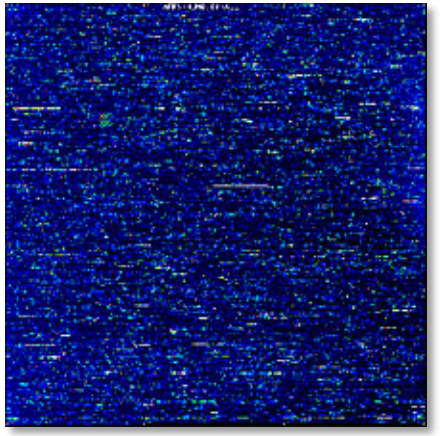

Functional Comparison

| | Correct Mechanistic Match | | |
|---|---|---|---|
| | 2/5 (40%) | 2/5 (40%) | 5/5 (100%) |

Karmaus and Martin, Unpublished

# Requirements and Potential Platforms for HT Transcriptomics





## Requirements

- Measure or infer transcriptional changes across the whole genome (or very close to it) (e.g. not subsets of 1000, 1500, 2500 genes)
- Compatible with 96- and 384-well plate formats (maybe 1536?) and laboratory automation
- Work directly with cell lysates (no separate RNA purification)
- Compatible with multiple cell types and culture conditions
- Low levels of technical variance and robust correlation with orthogonal measures of gene expression changes
- Low cost ($30 - $45 per sample or less)

## Potential Platforms

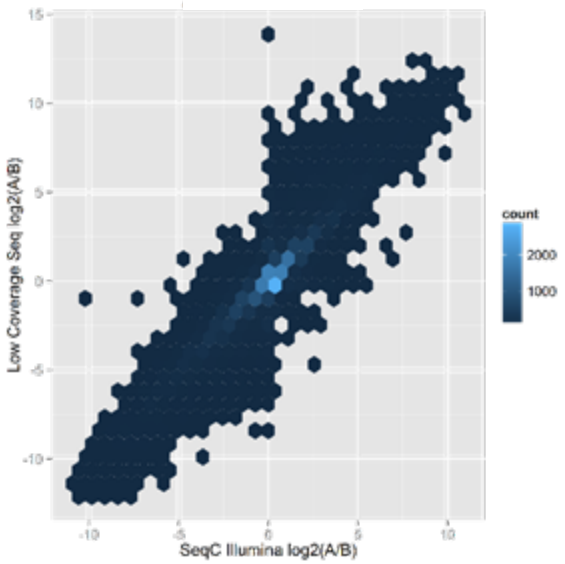
- Low coverage whole transcriptome RNA-seq (3 – 5 million mapped reads)
- Targeted RNA-seq (e.g., TempO-seq, TruSeq, SureSelect)
- Microarrays (e.g., Genechip HT)
- Bead-based (e.g., L1000)

# Technical Performance of the Three Sequencing Platforms

United States Environmental Protection Agency

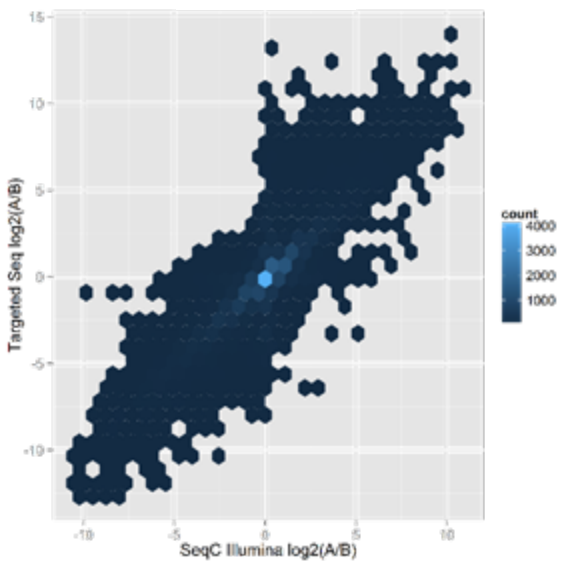
## Low Coverage

$r^2$ 0.83



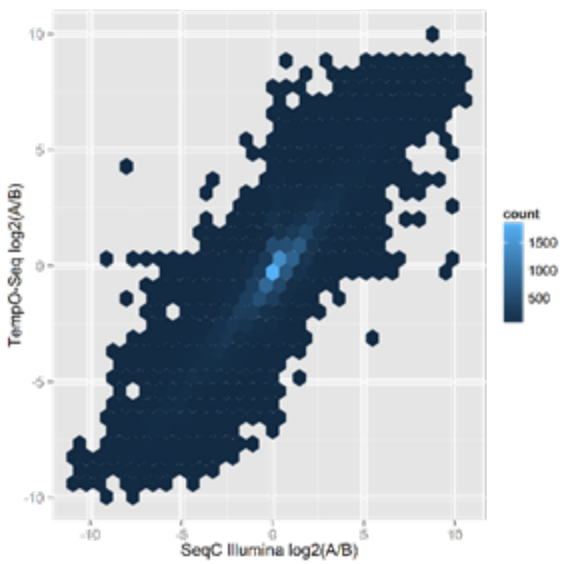

## TruSeq

$r^2$ 0.74




## TempO-Seq

$r^2$ 0.75




Data from MAQC II Samples

# HT Transcriptomics Next Steps

- Perform pilot study (Summer) to validate workflow and refine experimental design
- Initiate large scale screen (Fall/Winter)
    - Cell type: MCF7
    - Compounds: 1,000 (ToxCast Phase I/II)
    - Time Point: Single
    - Concentration Response: 8 (?)
- Perform secondary pilot study looking at cell type selection/ pooling strategies (Fall/Winter)
- Integrate HT transcriptomic platform with metabolic retrofit solution to allow screening +/- metabolism (FY17)
- **Explore partnerships to build community database of common chemical set across multiple cell types/lines**

# Other Ongoing Efforts

- **Curated chemical structure database of >1 million unique substances**
- Capability to retrofit high-throughput *in vitro* assays for metabolic competence
- **Software infrastructure to manage, use and share big data in toxicology**
- Methods to quantify uncertainty in all quantities
- **Read-across approaches that quantitatively include uncertainty**
- Pharmacokinetic models for hundreds of chemicals while understanding which chemical classes are well predicted and which ones have greater uncertainty
- High-throughput exposure models for thousands of chemicals with estimates of uncertainty
- **Non-targeted analytical measurements of chemical constituents in hundreds of consumer products**
- **Framework for streamlined validation of high-throughput *in vitro* assays**

EPA United States Environmental Protection Agency

# Challenges

- Technical limitations/obstacles associated with each technology (e.g., metabolism, volatiles, etc.)
- Moving from an apical to a molecular paradigm and defining adversity
- Predicting human safety vs. toxicity
- Combining new approaches to have adequate throughput and sufficiently capture higher levels of biological organization
- Systematically integrating multiple data streams from the new approaches in a risk-based, weight of evidence assessment
- Quantifying and incorporating uncertainty and variability
- Dealing with the validation
  - Defining a fit-for-purpose framework(s) that is time and resource efficient
  - Performance-based technology standards vs. traditional validation
  - Role of *in vivo* rodent studies and understanding their inherent uncertainty
- Legal defensibility of new methods and assessment products

# Acknowledgements


Stephanie Padilla
Tamara Tal
Eric Watt
Matt Martin
Rusty Thomas
Agnes Karmaus
Steve Simmons
Danica DeGroot
Keith Houck
John Wambaugh
Woody Setzer

NCCT Staff Scientists
Rusty Thomas
Kevin Crofton
Keith Houck
Ann Richard
Richard Judson
Tom Knudsen
Matt Martin
Grace Patlewicz
Woody Setzer
John Wambaugh
Tony Williams
Steve Simmons
Chris Grulke
Jeff Edwards

NCCT Postdocs
Todor Antonijevic
Audrey Bone
Swapnil Chavan
Danica DeGroot
Jeremy Fitzpatrick
Jason Harris
Dustin Kapraun
Max Leung
Kamel Mansouri
LyLy Pham
Prachi Pradeep
Eric Watt

NCCT Contractors
Nancy Baker
Dayne Filer
Parth Kothiya
Doris Smith
Jamey Vail
Sean Watford
Indira Thillainadarajah
Tommy Cathey



NIH/NCATS
Menghang Xia
Ruili Huang
Anton Simeonov

NTP
Warren Casey
Nicole Kleinstreuer
Mike Devito
Dan Zang