Computational Toxicology in EPA

Kevin M. Crofton
Deputy Director
National Center for Computational Toxicology

Central States SOT Meeting
18 November 2016
Iowa City

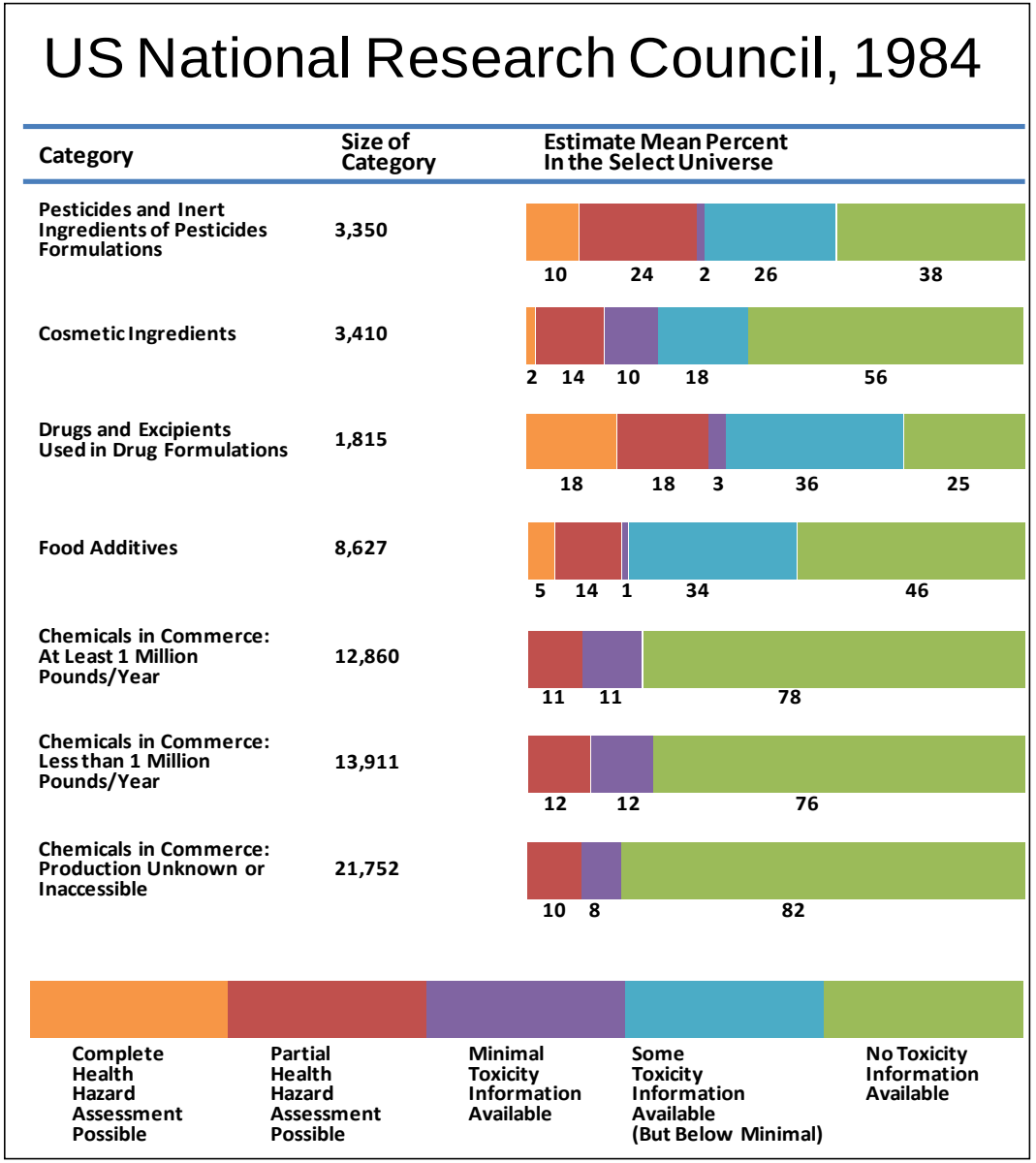
Outline

- **The Problem**
  - Thousands and thousands of chemicals with no hazard or exposure data
- **Small Segue**
  - The Importance of Matching Data Type to the Decision Context – "fit for purpose"
- **Demonstrating Progress**
  - Chemistry
  - ToxCast and Tox21 - Bioactivity
  - Exposure Predictions
  - Putting it all Together
- **Challenges (some of them….)**

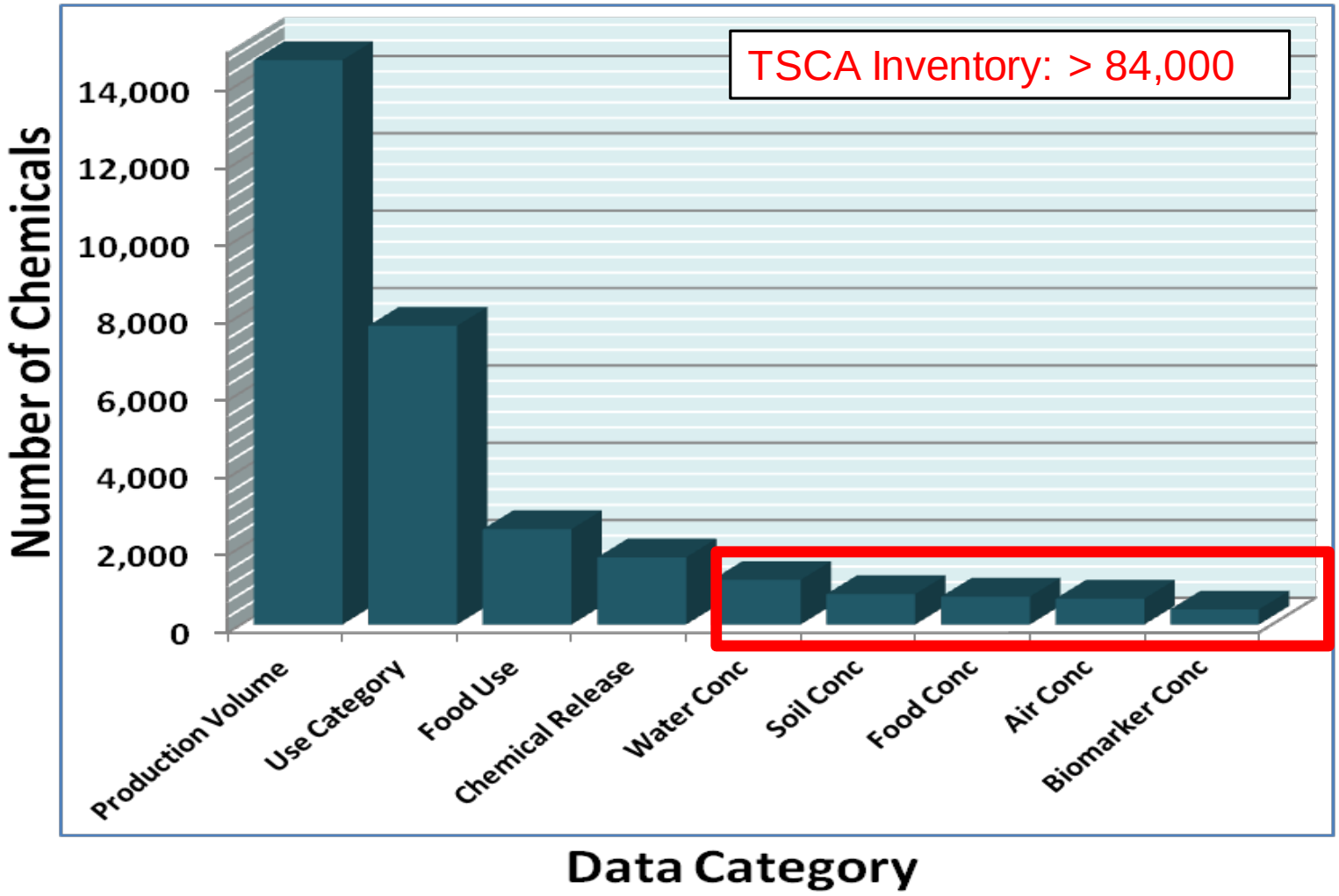
Problem: The Chemical Hazard Universe

1974 US NRC report

- Major challenge is too many chemicals and not enough data
- Estimated number of chemicals = 65,725
- Number of chemical with no toxicity data of any kind = ~46,000



Problem: The Chemical Exposure Universe

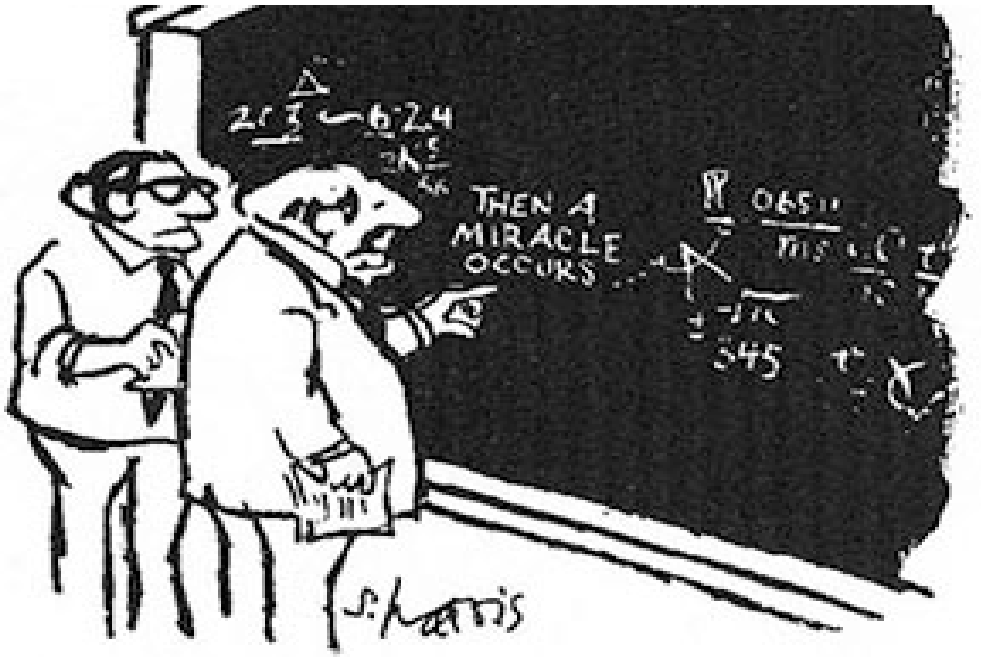


P.P. Egeghy et al. *Sci Total Environ.* 414 (2012) 159–166

Segue - Matching Data Type and Uncertainties to Decision Context

It is critical to understand the uncertainties in the data

And match them to the regulatory decision context



"I think you should be more explicit here in step two."

EPA Office	Assessment "Workflows"	Historical Throughput	
OPPTS	Premanufacture Notice (PMN) New chemicals Significant New Use Rule (SNUR) Existing chemicals	~1000/yr 90d/chem ~84,000 total	
	Current Chemical Risk *(new program)*	~10 total	I
	DFE / Green Chemistry	~2500	I, II, III
OSCP	Endocrine Screening Program	~10-20/year	
OPP	Pesticide registration (PR)	~10 new/yr ~5	I
	Pesticide re-registration	~ 24,	
OW	Chemical Contaminant List	~6,000 total	
	Regulatory Actions on CCL	6yr 90 total	I
	Unregulated Contaminant Monitoring	30/5yr	I
	Drinking Water Health Advisories (MCLs)	~80 total	II, III
ORD NCEA	IRIS	~3/yr ~540 total	I
	PPRTV	400-500	II,III

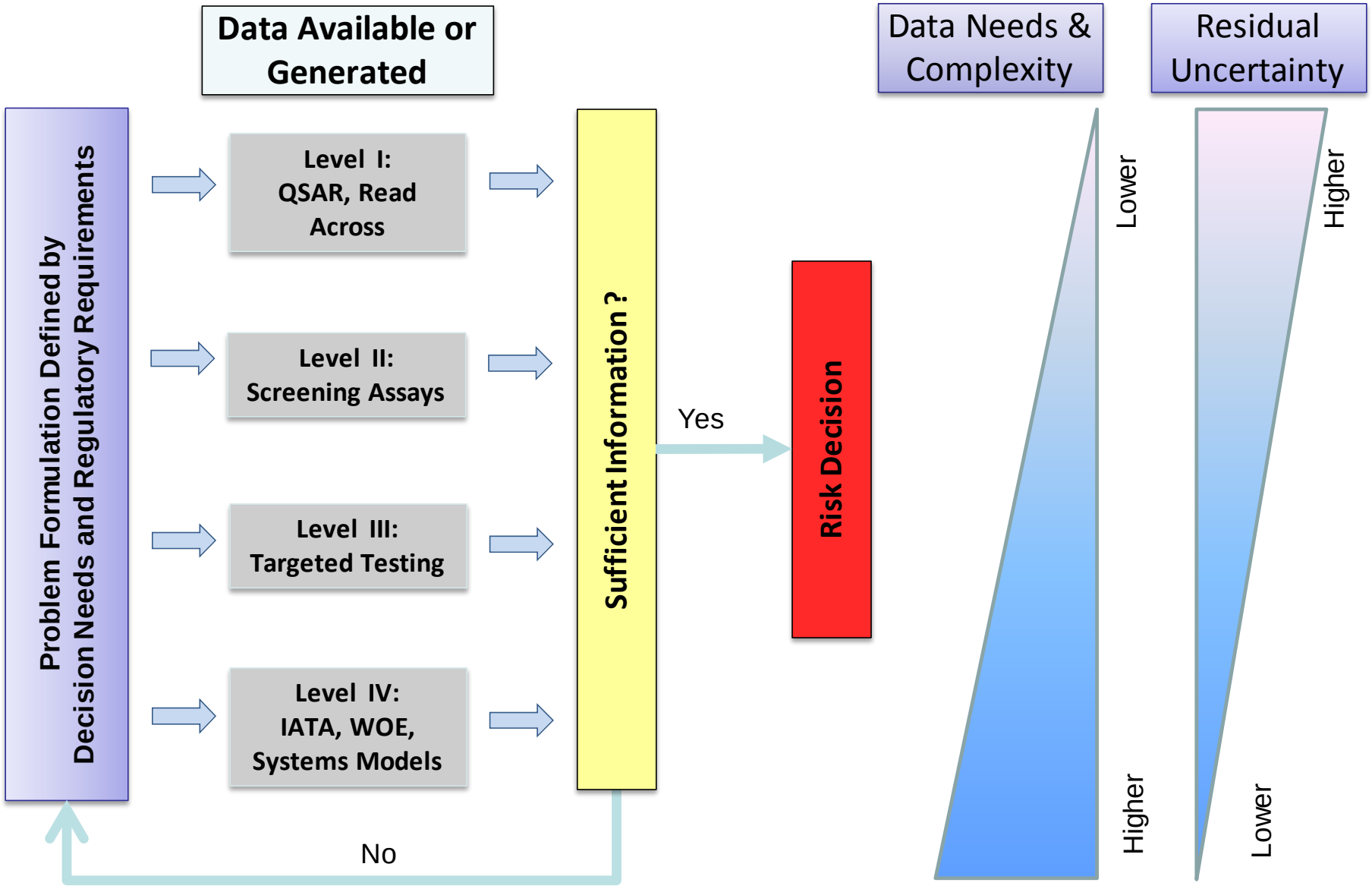
~1000/year
90 days/chemical
Limited data., eg structure, LogP

~10/yr
$ millions of data

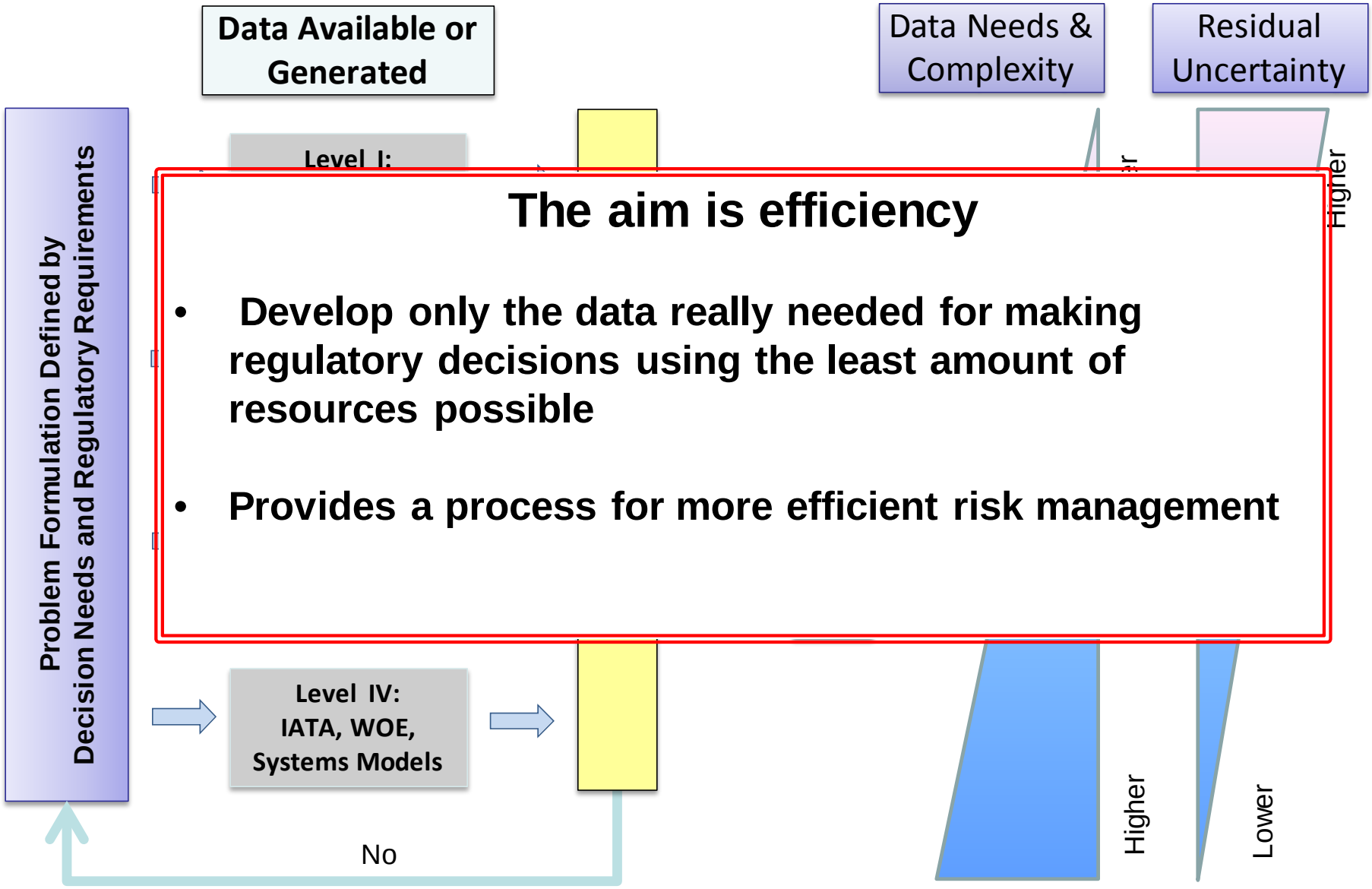
I. Data rich – Extensive guideline studies
II. Data partial – Some acute in vivo and in vitro data, SAR and exposure modeling
III. Data minimal to none – only chemical structure, SAR and exposure modeling

Courtesy of I. Shah

Matching Data Type and Uncertainties to Decision Context

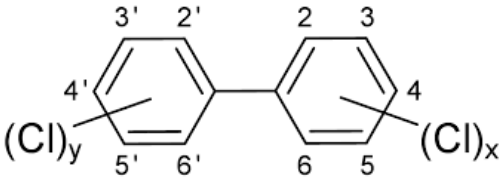


Matching Data Type and Uncertainties to Decision Context

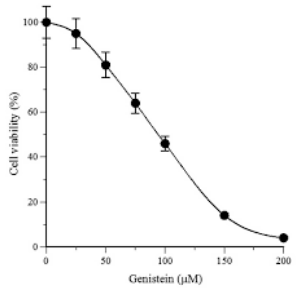


The Beginning of NCCT

What is Necessary to Begin Solving the Problem of Too Many Chemicals With No Exposure or Hazard Information



1. **Chemical curation**
   - Everything starts with chemical structure

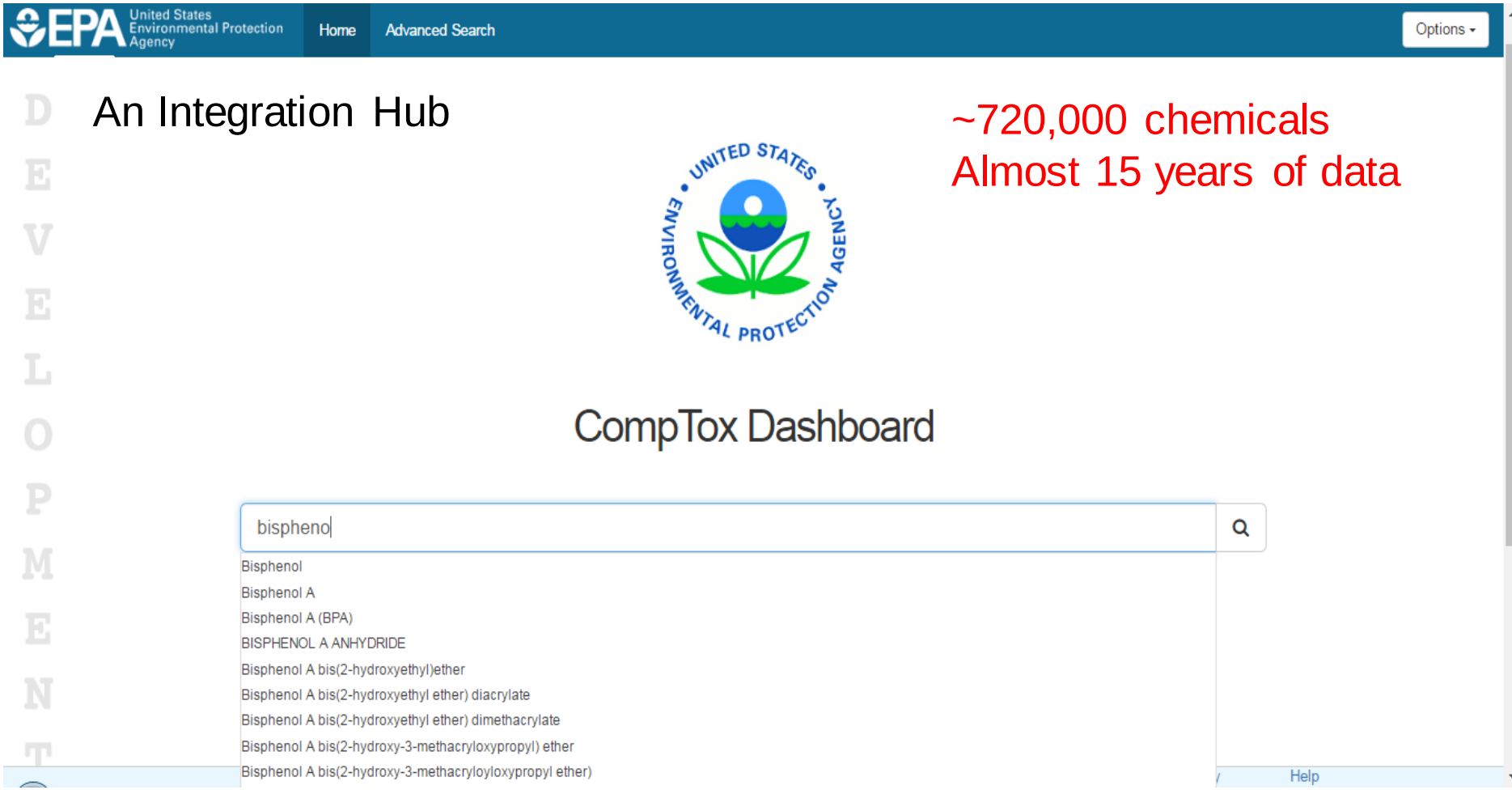


2. **Prediction of hazard (or bioactivity)**
   - Need fast efficient testing methods

3. **Predictions of exposure**
   - Need new models that predict or measure exposures

4. **Putting it all together**
   - Models that integrate this into estimates of risk
   - Tools that can be used by risk managers

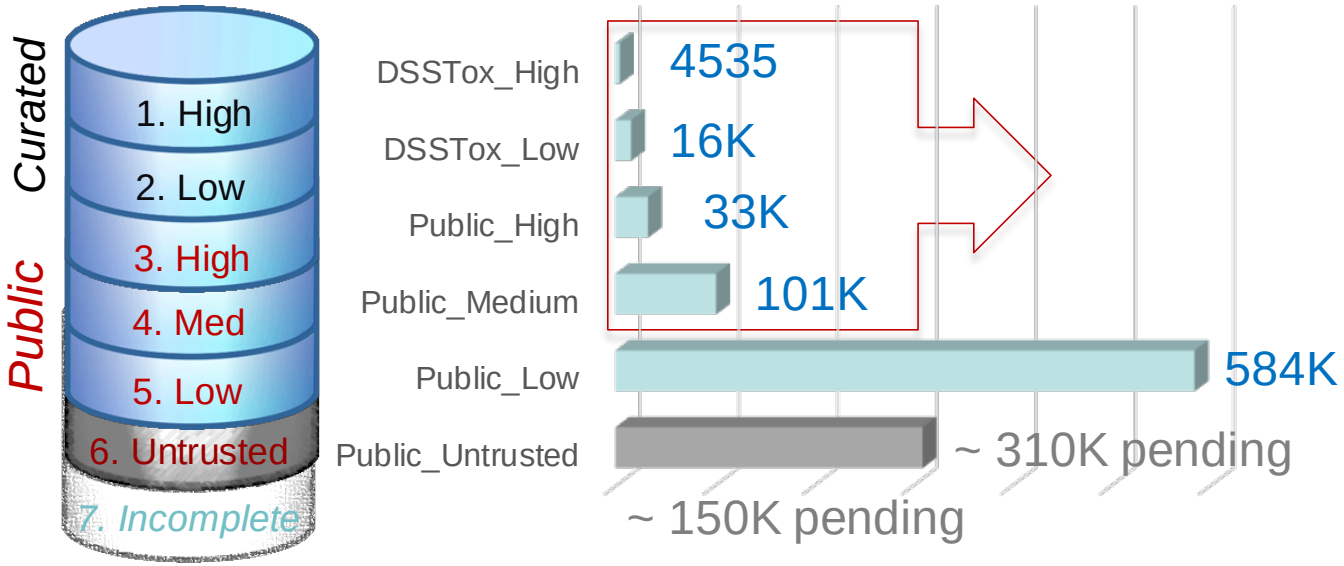
CompTox Dashboard - Chemistry



https://comptox.epa.gov

Even in Chemistry there is Uncertainty

720,000 chemicals ranked on confidence

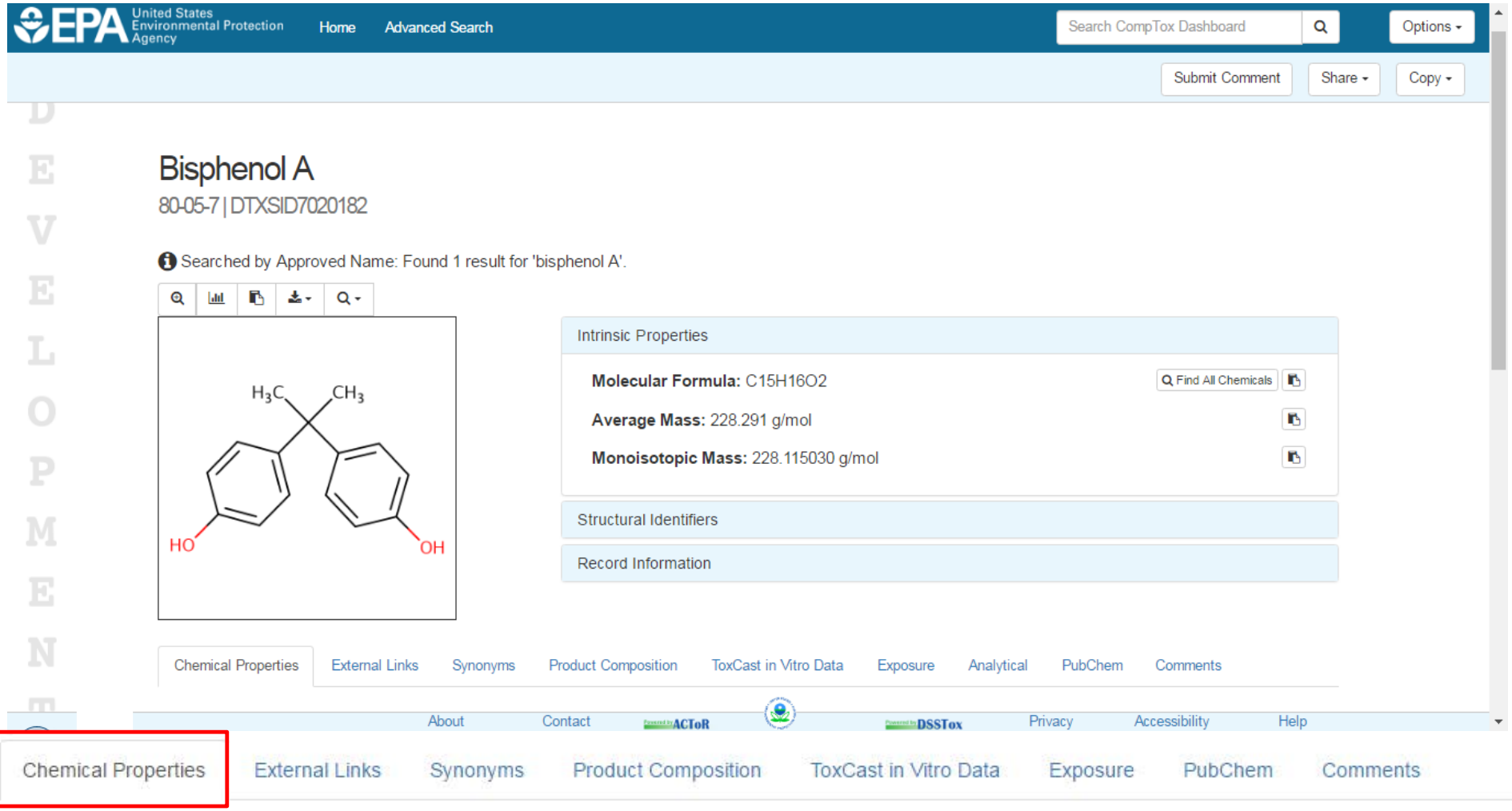


QC Levels

DSSTox_High:	Hand curated and validated
DSSTox_Low:	Hand curated and confirmed using multiple public sources
Public_High:	Extracted from EPA SRS and confirmed to have no conflicts in ChemID and PubChem
Public_Medium:	Extracted from ChemID and confirmed to have no conflicts in PubChem
Public_Low:	Extracted from ACToR or PubChem
Public_Untrusted:	Postulated, but found to have conflicts in public sources

Richard et al. Chem Res Toxicol. 29:1225-1251, 2016

Dashboard Example Bisphenol A



Generating Bioactivity Data

ToxCast and Tox21 Programs

- ToxCast – EPA program
  - Multi-year research program started in 2007
  - Use automated in vitro chemical screening technologies to expose living cells or isolated proteins to chemicals where changes in biological activity may suggest potential toxic effects
  - Chemical library
    - ~3400 environmentally relevant chemicals

      http://www.epa.gov/ncct/toxcast/
- Tox21 – Collaborative project
  - US EPA, NIH/NCATS, NIH/NIEHS/NTP and FDA
  - aimed at developing better toxicity assessment methods using HTS.
  - Chemical library
    - ~8,500 chemicals, including environmental chemicals, food additives and pharmaceuticals
      http://www.ncats.nih.gov/research/reengineering/tox21/tox21.html

Increased Throughput Required Shift to Molecular/Pathway Approaches

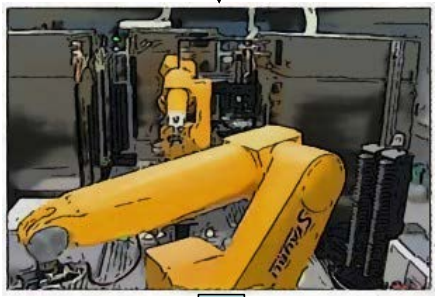
ToxCast	Tox21

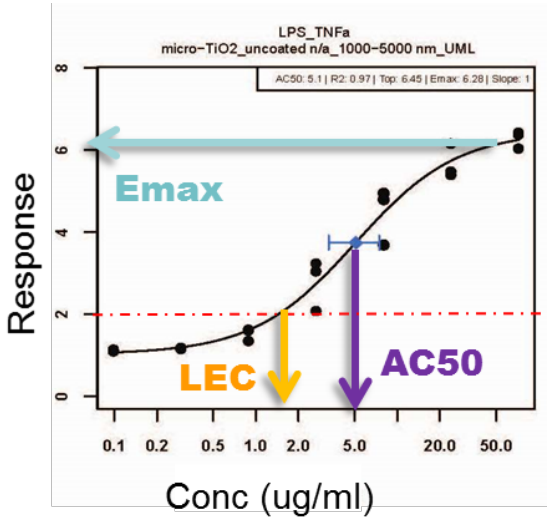
~600 Cell & biochemical assays

~1,000 Chemicals

~30 Cell & biochemical assays

~8,000 Chemicals

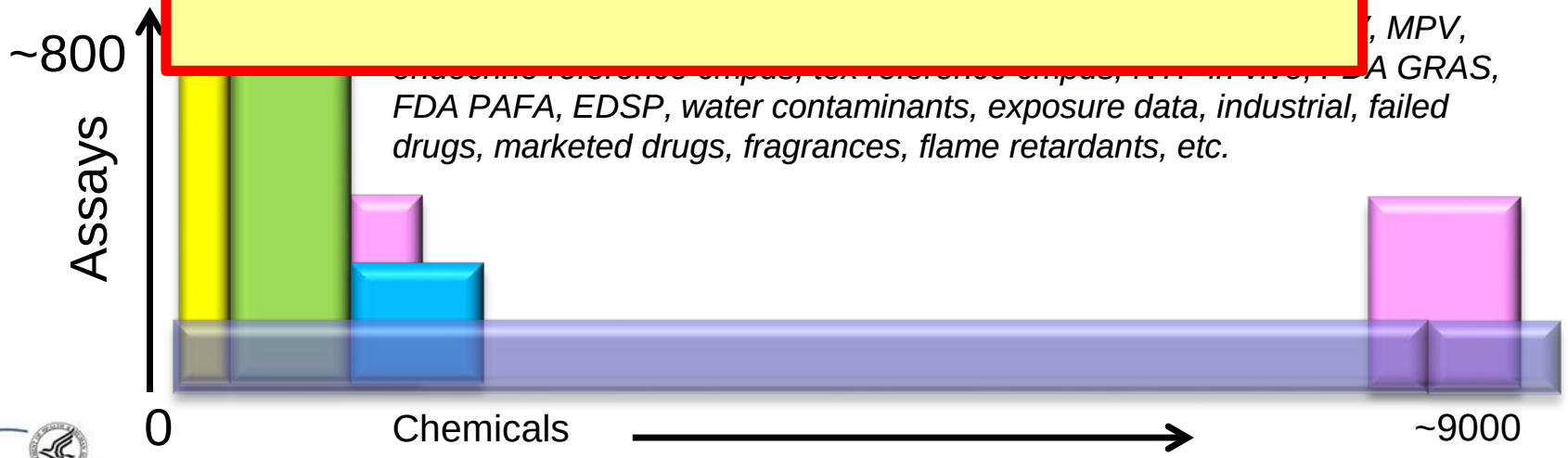




ToxCast & Tox21: Chemicals, Data and Release Timelines

Set	Chemicals	Assays	Endpoints	Completion	Available
ToxCast Phase I	293	~600	~700	2011	Now
ToxCast Phase II	767	~600	~700	03/2013	Now
ToxCast E1K					Now
Tox21					Ongoing
ToxCast Pha					2016

72 million data points
2.8 million conc response curves

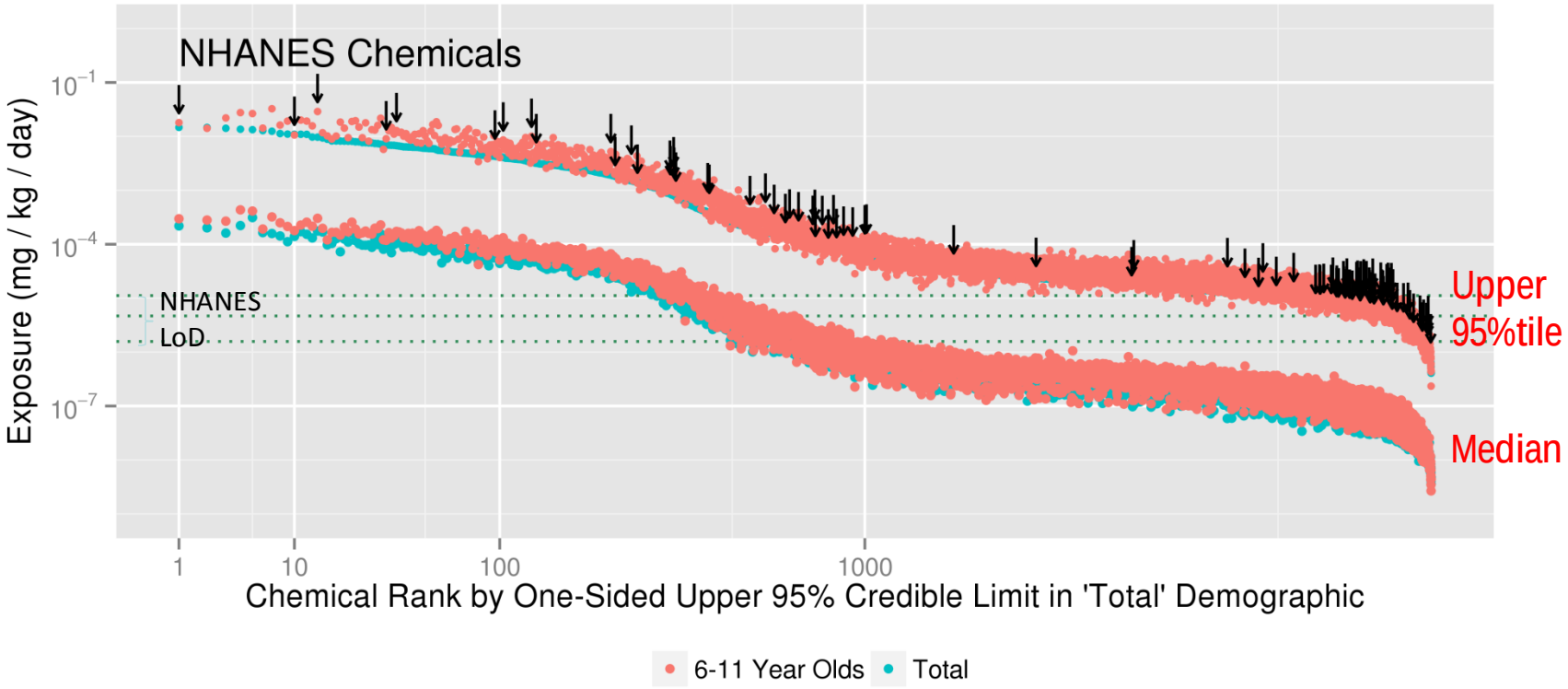


ExpoCast
HTP Exposure Predictions

- **For years exposure science has lagged behind**
  - Most models require extensive information on production, use, fate and transport and rely on empirical data *(no measurement = no exposure?)*

- **ExpoCast**
  - Exposure predictions based on:
    - pChem, production values, fate and transport, and product use categories (e.g., industrial, pesticide use, consumer personal care)
    - Industrial vs consumer use
  - Yields exposure estimates and Baysian confidence

Exposure Predictions for 7968 Chemicals & Comparison to NHANES

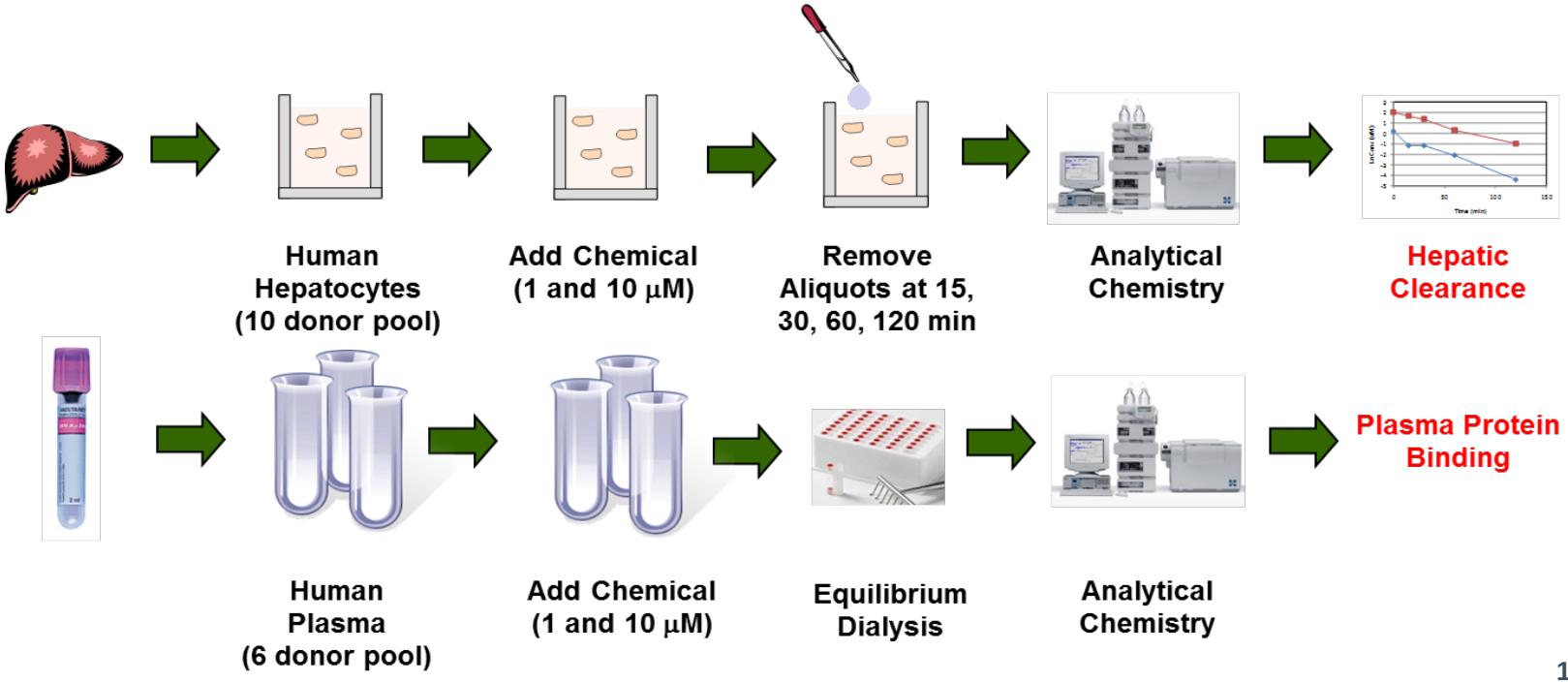


- NHANES – US National Study – measures exposures in human serum and urine
- Chemicals currently monitored by NHANES are distributed throughput the predictions
- Shows accuracy of the prediction model

***Wambaugh et al Environ. Sci. Technol.*, 2014, *48* (21), pp 12760–12767**

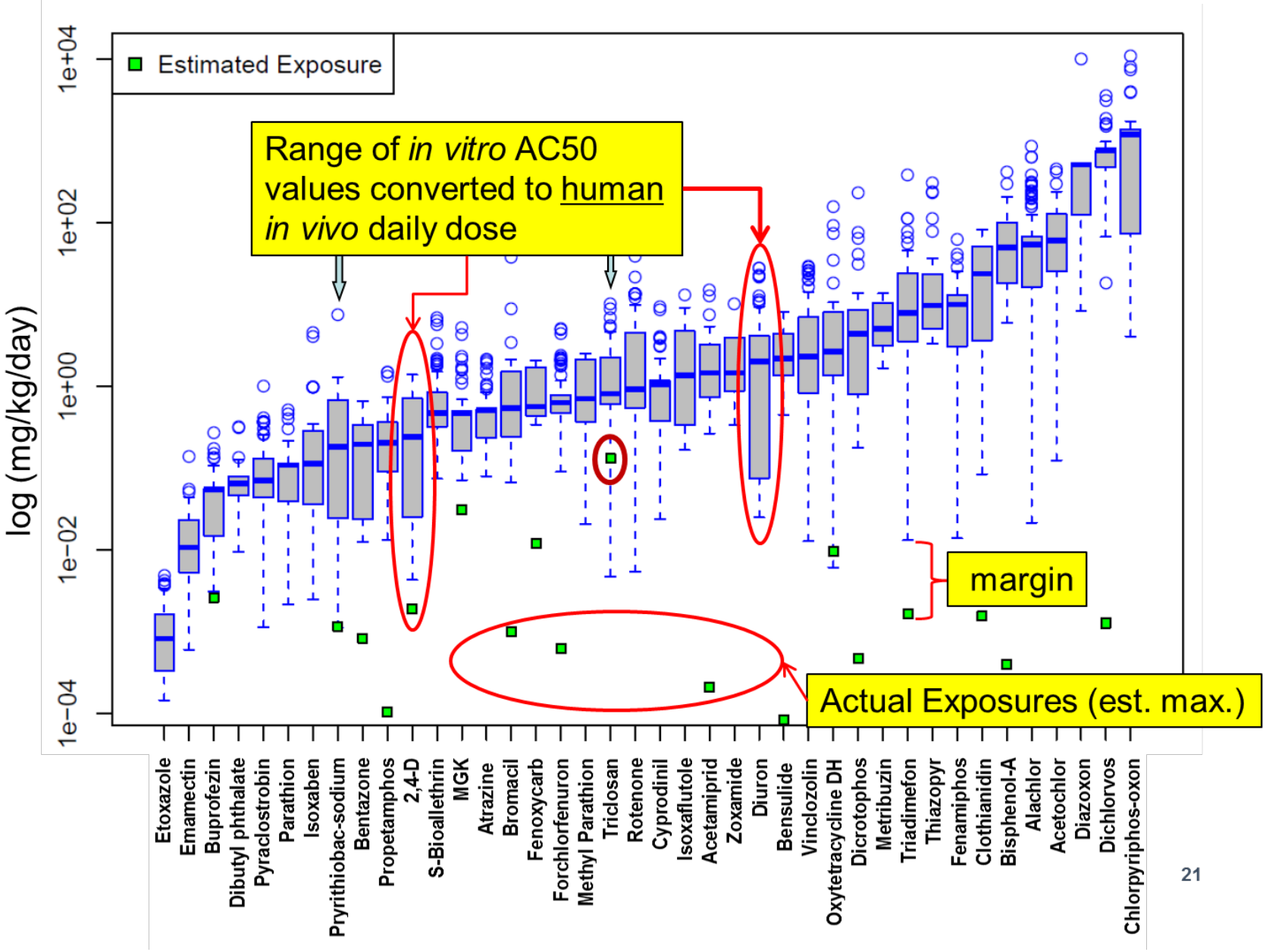
Estimating Daily Dose with Reverse ToxicoKinetics (rTK)

- VERY PK models – *measure only 2 parameters*
  - in vitro hepatic clearance disappearance of parent compound
  - serum protein binding values
- Provides scaling from concentration in which there is in vitro biological activity to in vivo activity dose (mg/kg/day)



(Rotroff et al, ToxSci 2010, Wetmore et al, ToxSci 2012)

Estimating Exposures for *in vitro bio*activity



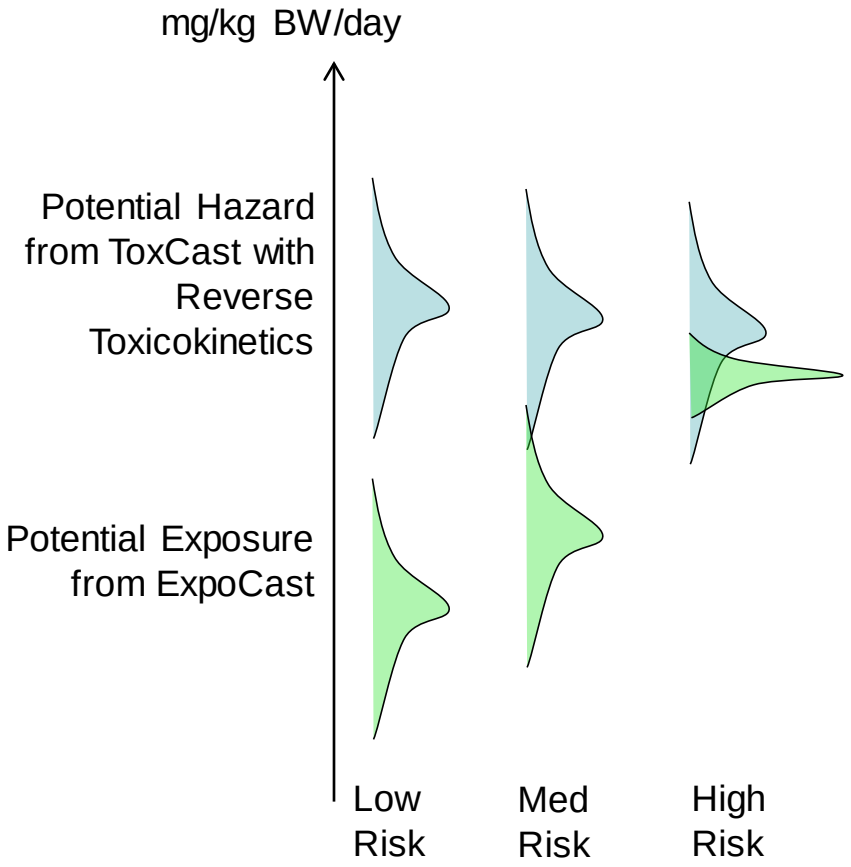
Wetmore et al Tox Sci 2012

Putting It All Together
HT Prioritization

Risk is the product of hazard and exposure

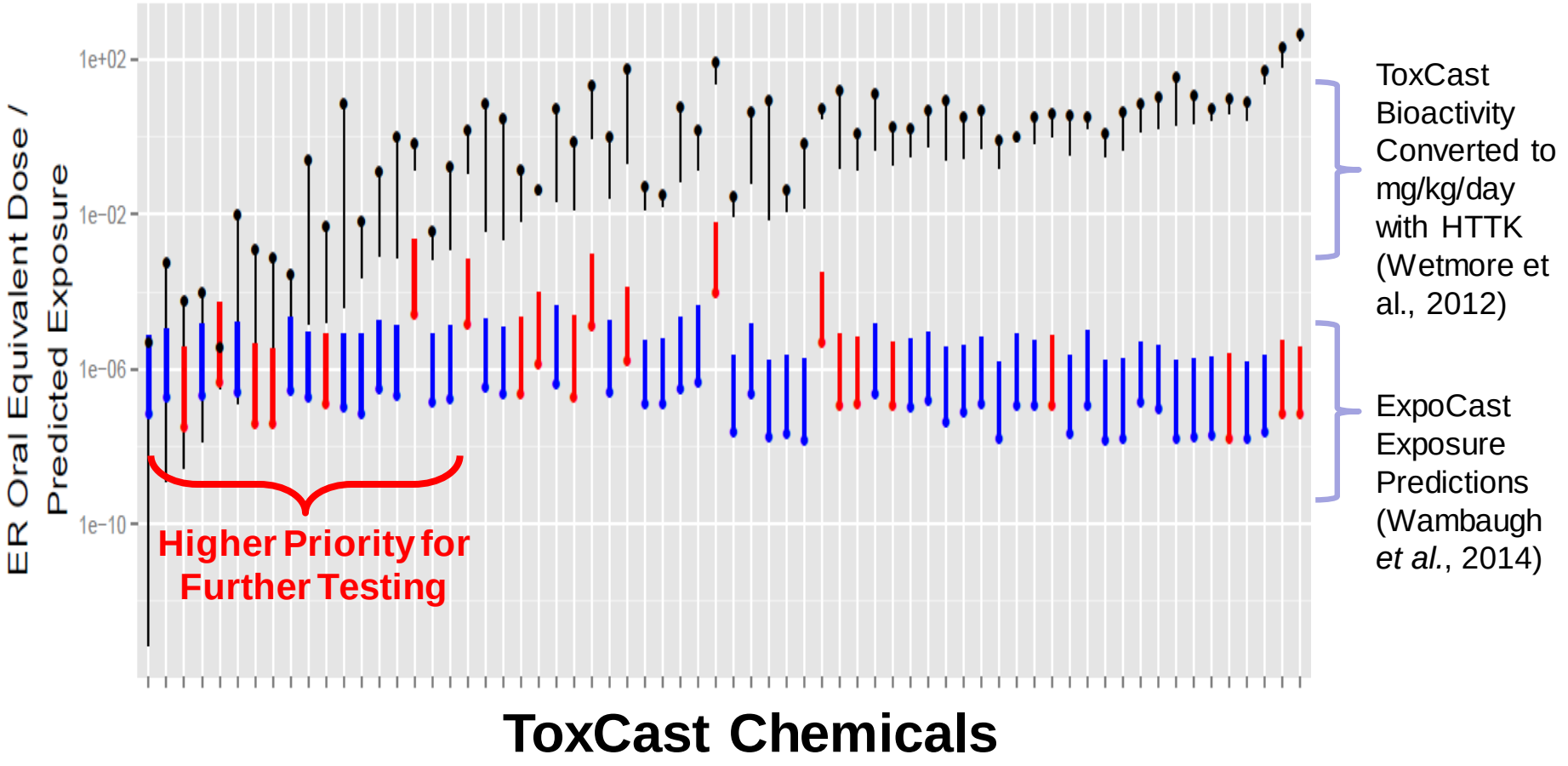
Use rTK convert bioactive concentrations to daily dose

Combine with exposure prediction



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

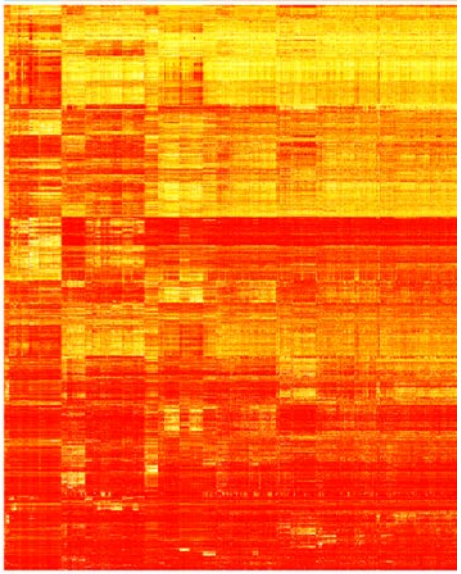
Combining Bioactivity-Base Dose and Exposure Estrogen Active Chemicals



Prioritization = test the chemicals that might be the worst, first!

Broad Success using High-Throughput Screening Approaches

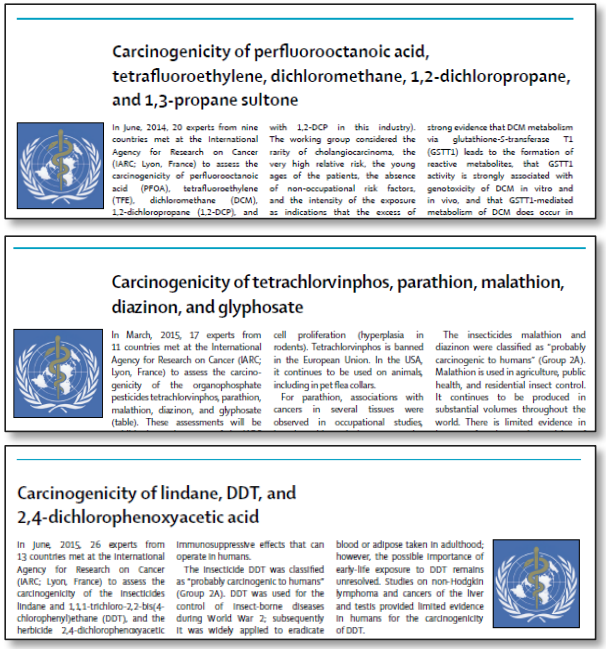
Group Chemicals by Similar Bioactivity For Predictive Modeling



Predictive Models:

- Developmental Toxicity
- Vasculogenesis
- Reproductive Toxicity

Provide Mechanistic Support for Hazard ID



Carcinogenicity of perfluorooctanoic acid, tetrafluoroethylene, dichloromethane, 1,2-dichloropropane, and 1,3-propane sultone

In June, 2014, 20 experts from nine countries met at the International Agency for Research on Cancer (IARC; Lyon, France) to assess the carcinogenicity of perfluorooctanoic acid (PFOA), tetrafluoroethylene (TFE), dichloromethane (DCM), 1,2-dichloropropane (1,2-DCP), and with 1,2-DCP in this industry). The working group considered the rarity of cholangiocarcinoma, the very high relative risk, the young ages of the patients, the absence of non-occupational risk factors, and the intensity of the exposure as indications that the excess of strong evidence that DCM metabolism via glutathione-S-transferase T1 (GSTT1) leads to the formation of reactive metabolites, that GSTT1 activity is strongly associated with genotoxicity of DCM in vitro and in vivo, and that GSTT1-mediated metabolism of DCM does occur in

Carcinogenicity of tetrachlorvinphos, parathion, malathion, diazinon, and glyphosate

In March, 2015, 17 experts from 11 countries met at the International Agency for Research on Cancer (IARC; Lyon, France) to assess the carcinogenicity of the organophosphate pesticides tetrachlorvinphos, parathion, malathion, diazinon, and glyphosate (table). These assessments will be cell proliferation (hyperplasia in rodents). Tetrachlorvinphos is banned in the European Union. In the USA, it continues to be used on animals, including in pet flea collars.

For parathion, associations with cancers in several tissues were observed in occupational studies, The insecticides malathion and diazinon were classified as "probably carcinogenic to humans" (Group 2A). Malathion is used in agriculture, public health, and residential insect control. It continues to be produced in substantial volumes throughout the world. There is limited evidence in

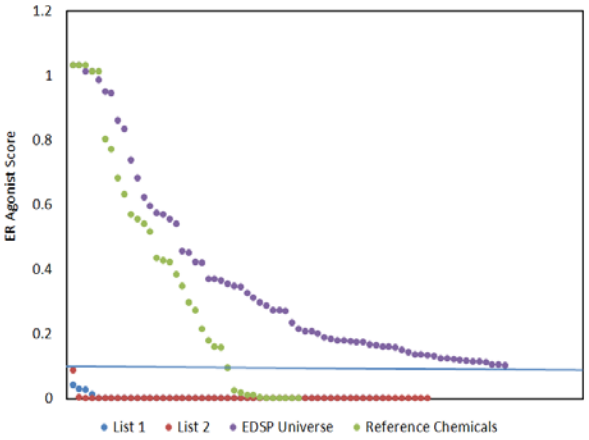
Carcinogenicity of lindane, DDT, and 2,4-dichlorophenoxyacetic acid

In June, 2015, 26 experts from 13 countries met at the International Agency for Research on Cancer (IARC; Lyon, France) to assess the carcinogenicity of the insecticides lindane and 1,1,1-trichloro-2,2-bis(4-chlorophenyl)ethane (DDT), and the herbicide 2,4-dichlorophenoxyacetic immunosuppressive effects that can operate in humans.

The insecticide DDT was classified as "probably carcinogenic to humans" (Group 2A). DDT was used for the control of insect-borne diseases during World War 2; subsequently it was widely applied to eradicate blood or adipose taken in adulthood; however, the possible importance of early-life exposure to DDT remains unresolved. Studies on non-Hodgkin lymphoma and cancers of the liver and testis provided limited evidence in humans for the carcinogenicity of DDT.

ToxCast Data used in WOE decisions by International Agency for Researh on Cancer (IARC Monographs 110, 112, 113)

Prioritization of Chemicals Based on Potency for Further Testing



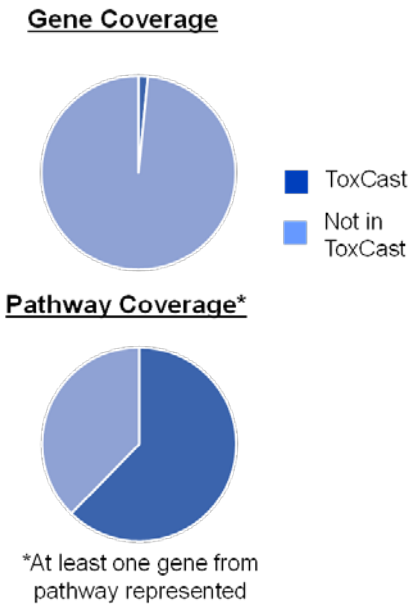
FIFRA SAP, Endocrine Disruption Screening Program Dec 2014

Prioritization

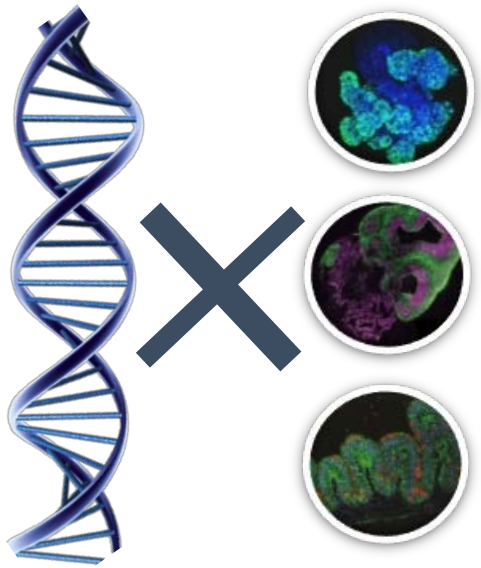
And for first time replacement!

Challenges: Increasing Biological Coverage

Gene coverage in ToxCast is low = ~300



Ongoing pilot project using **cost-efficient** whole gene sequencing platforms



Whole Transcriptome Technologies

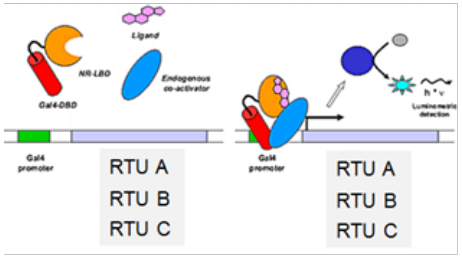
Diverse Cell Types & Culture Systems

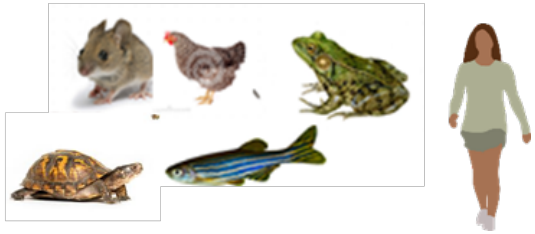
Goal: use multiple cell types to cover entire genome

Challenges: Increasing Species Coverage

- ToxCast & Tox21 are mostly human based
- Pilot project using Attagene system – insert receptor ligand binding domains from multiple species
- Multiple readouts of nuclear receptor hits from one cell

Multispecies Attagene *Trans* Reporter Assay



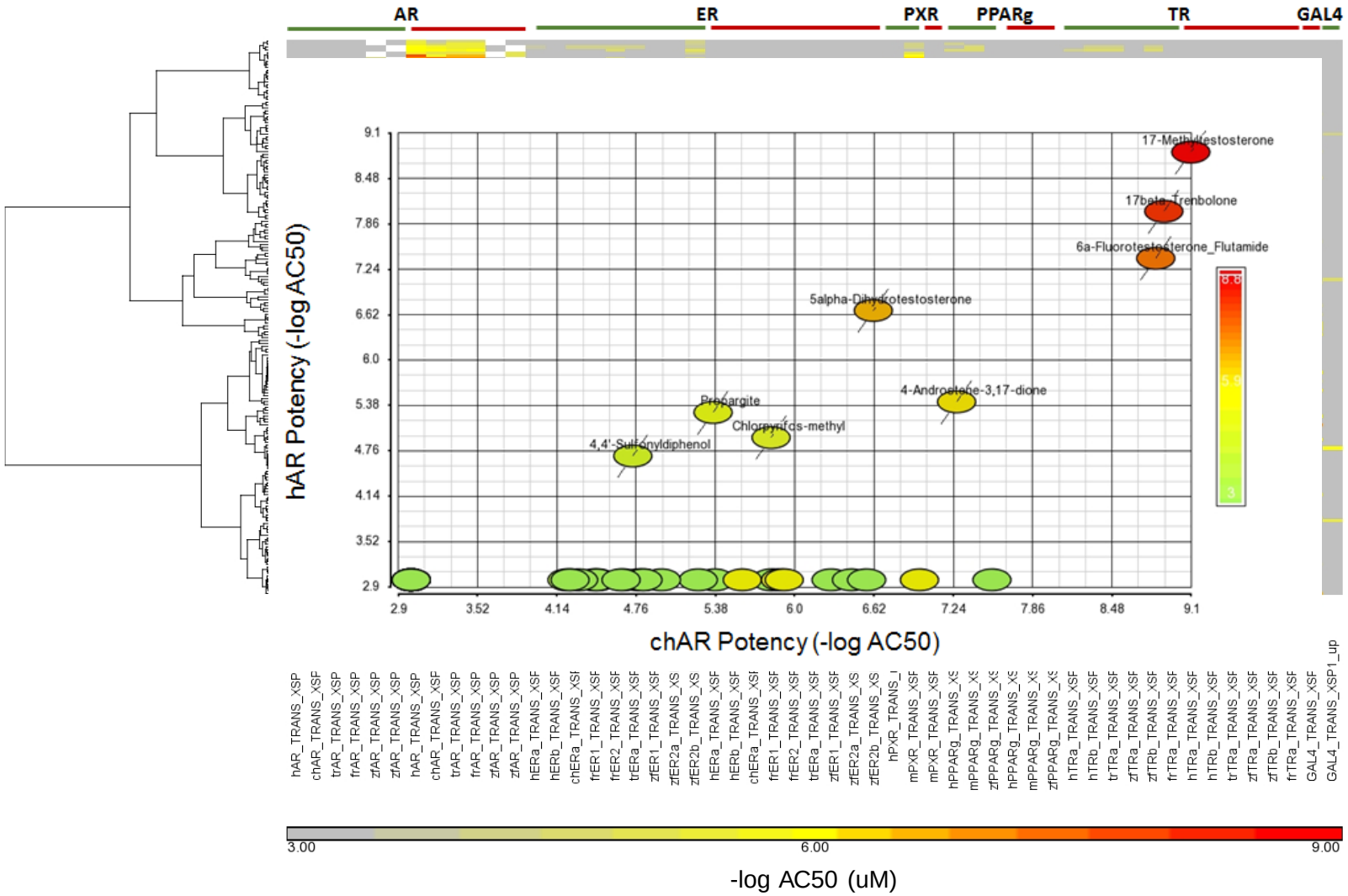


Receptor Family	Receptor Name	Species
Estrogen Receptor	ERa	Human
Estrogen Receptor	ERb	Human
Estrogen Receptor	ER1	Zebrafish
Estrogen Receptor	ER2a	Zebrafish
Estrogen Receptor	ER2b	Zebrafish
Estrogen Receptor	ERa	Chicken
Estrogen Receptor	ER1	Frog
Estrogen Receptor	ER2	Frog
Estrogen Receptor	ERa	Turtle
Estrogen Receptor	AR	Human
Estrogen Receptor	AR	Chicken
Estrogen Receptor	AR	Turtle
Estrogen Receptor	AR	Frog
Estrogen Receptor	AR	Zebrafish
Peroxisome Proliferator Activated Receptor γ	PPARg	Mouse
Peroxisome Proliferator Activated Receptor γ	PPARg	Zebrafish
Peroxisome Proliferator Activated Receptor γ	PPARg	Human
Pregnane X Receptor	PXR	Mouse
Thyroid Receptor	TRa	Turtle
Thyroid Receptor	TRb	Zebrafish
Thyroid Receptor	TRb	Zebrafish
Thyroid Receptor	TRa	Frog
Thyroid Receptor	TRa	Human
Thyroid Receptor	TRb	Human
Controls	M-06	NA
Controls	GAL4	NA
Controls	M-19	NA
Controls	m-32	NA
Controls	m-61	NA

- Host cell: human HepG2
- 100 chemicals with ER, AR, TR, PPAR activity tested in concentration-response
- Pilot data using positive and negative reference chemicals is promising

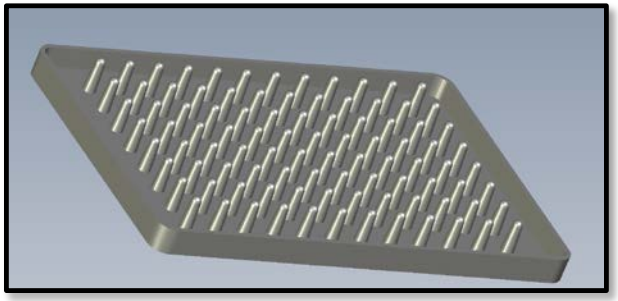
Houck, unpublished

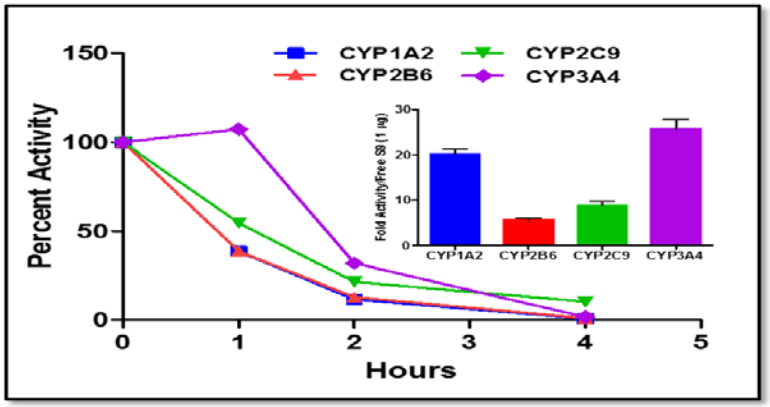
Pilot Data: Concordance and Cross-Species Differences in Nuclear Receptor Responses



Houck, unpublished

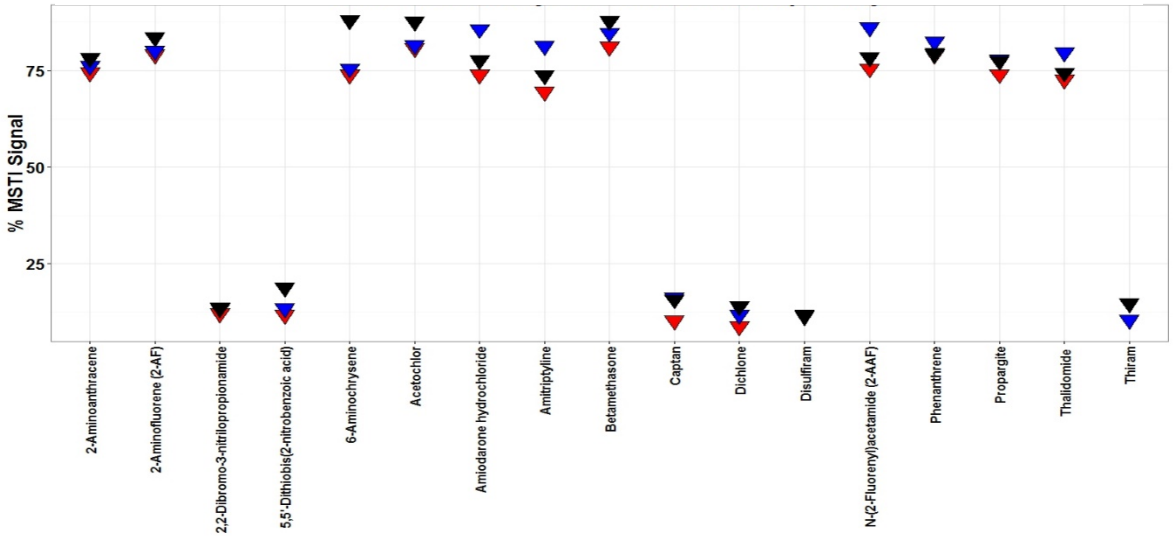
Challenges: Retrofitting Assays with Metabolic Activity





Cyp activity over time of encapsulated S9 fraction

MSTI Assay - An increase in electrophilicity was detected as a decrease in the fluorescent signal.



DeGroot and Simmons, Unpublished

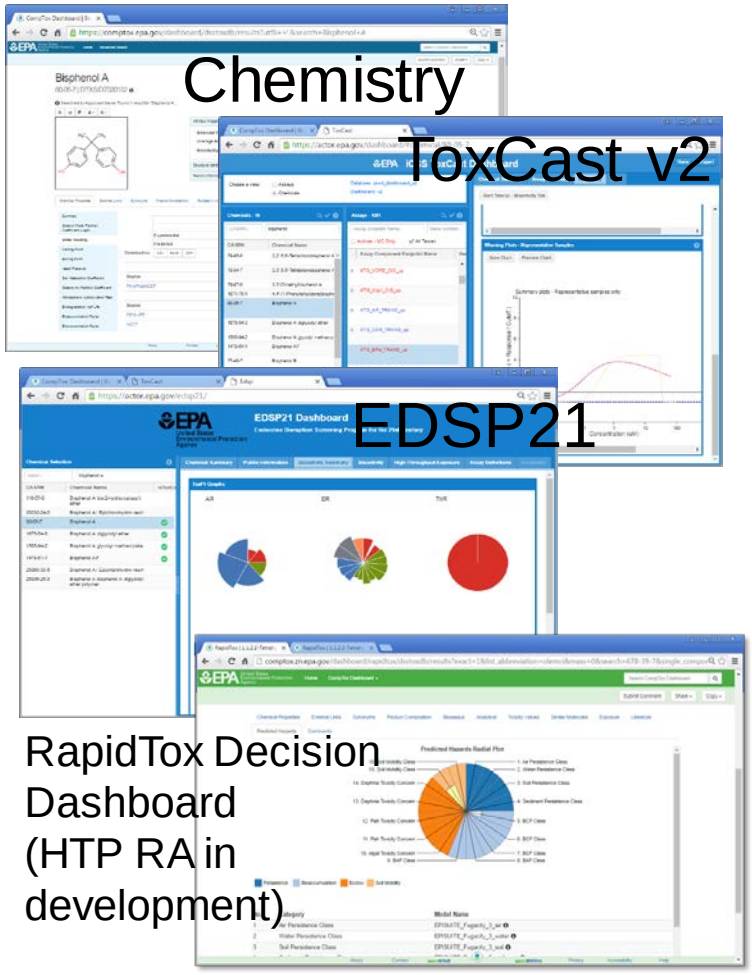
Challenges: Unless data is available and useful it will not be used

2011 Initial ToxCast Phase I Data Delivered as Flat Files

You need a bioinformatics degree for this to useful

2013 Dashboard with Limited Search, Visualization, and Export Functionality

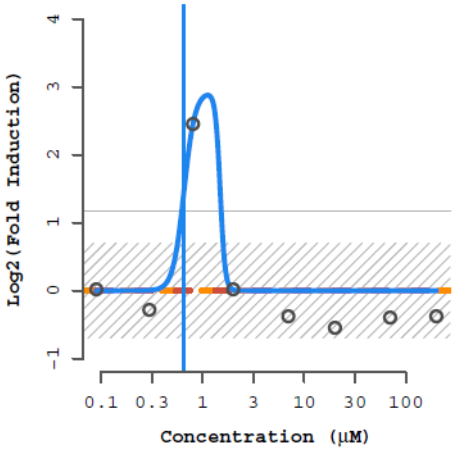
"Better, but still difficult to really get what you want without help from NCCT"



RapidTox Decision Dashboard (HTP RA in development)

Building data, analyses and visualization tools that allow for more rapid development of specific Decision Support Dashboards

Challenges: Regulatory Applications Require Transparency

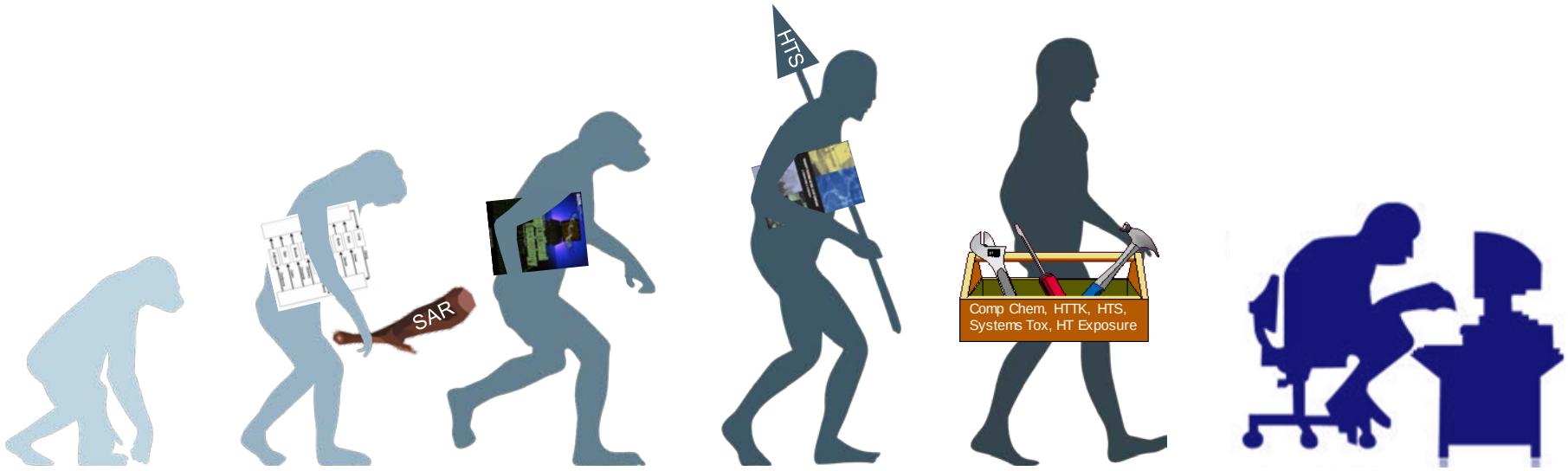



```
ASSAY:    AEID117 (ATG_ERa_TRANS)

NAME:     Thioglycolic acid
CHID:     26141    CASRN: 68-11-1
SPID(S): TX007664
L4ID:     420385

HILL MODEL (in red):
          tp        ga        gw
val:   3.1e-11   -2.15     0.416
sd:    NaN       NaN       NaN

GAIN-LOSS MODEL (in blue):
          tp        ga        gw        la        lw
val:   2.93      -0.184    8         0.173     18
sd:    3.56      0.334     9.48      5.82      814

          CNST         HILL         GNLS
AIC:   20.14        26.14        17.79
PROB:  0.23         0.01         0.76
RMSE:  0.92         0.92         0.32

MAX_MEAN: 2.45       MAX_MED: 2.45       BMAD: 0.233

COFF: 1.17    HIT-CALL: 1    FITC: 50    ACTP: 0.77
```

```
FLAGS:
Only one conc above baseline, active
Borderline active
```

- **Public release policy of from Tox21 and ToxCast Programs** (raw & processed data, all publications, all processing and modeling code)
- **Transparent ToxCast data analysis pipeline**
  - Data quality flags to indicate concerns with chemical purity and identity, noisy data, and systematic assay errors
  - Publicly available as an R package
- **Tox21 and ToxCast chemical libraries have undergone analytical QC and results publicly available**
- **Public posting of ToxCast procedures**
  - Chemical Procurement and QC
  - Data Analysis
  - Assay Characteristics and Performance
- **Recently completed external audit** on ToxCast data and data analysis pipeline

Next Phase… Evolution Towards a Truly Predictive Science



c.1500 – 1960s | c. 1970s - 2002 | 2003 - 2006 | 2007 - 2010 | 2011 - 2017 | 2017 - ?

Acknowledgements

Tox21 Colleagues:
- NTP Crew
- NCATS Collaborators
- FDA Collaborators

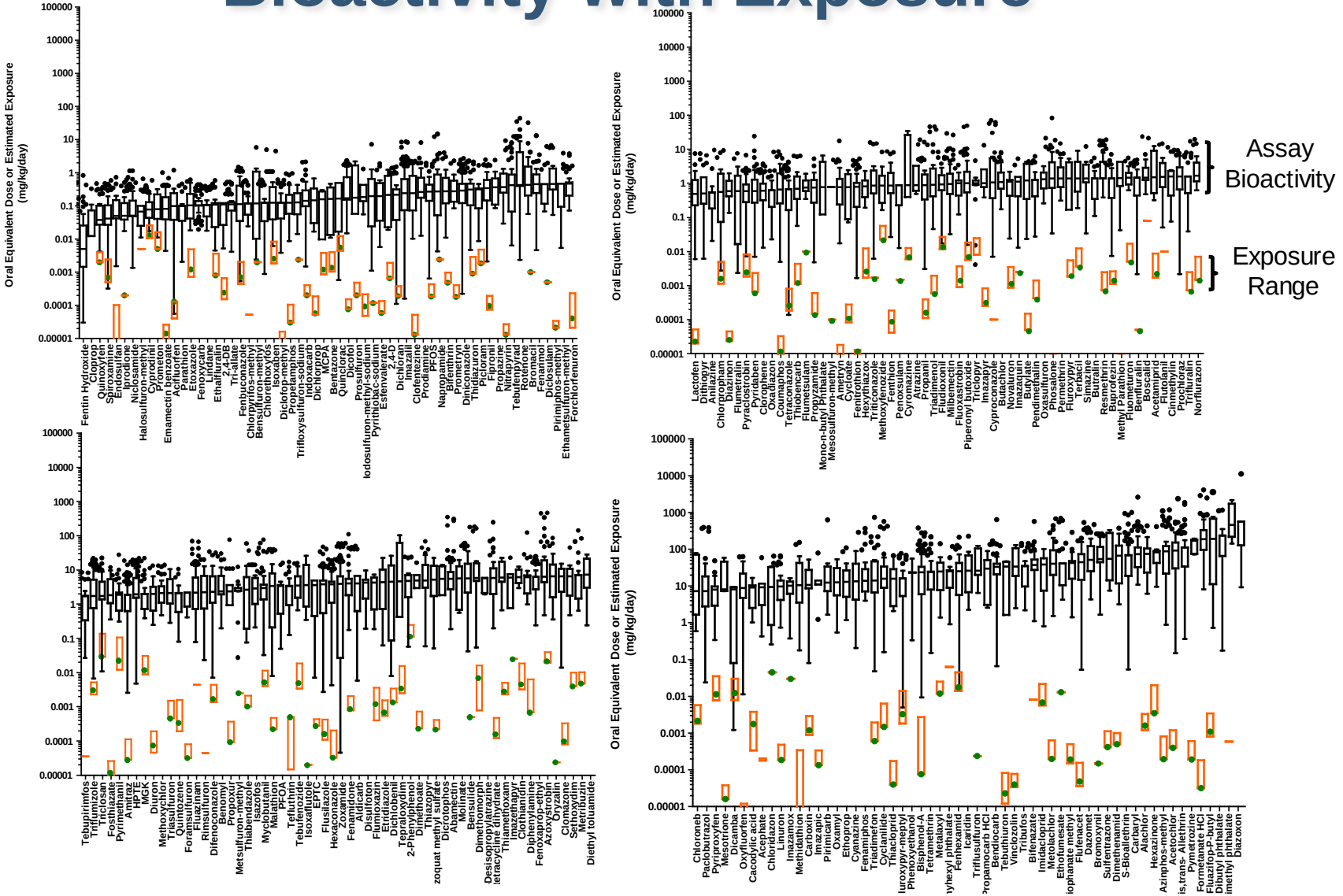
EPA Colleagues:
- NERL
- NHEERL
- NCEA
- EPA Program Offices

External Stakeholders
- Health Canada
- CalEPA
- EDF
- EU Joint Research Center
- ECHA

EPA's National Center for Computational Toxicology

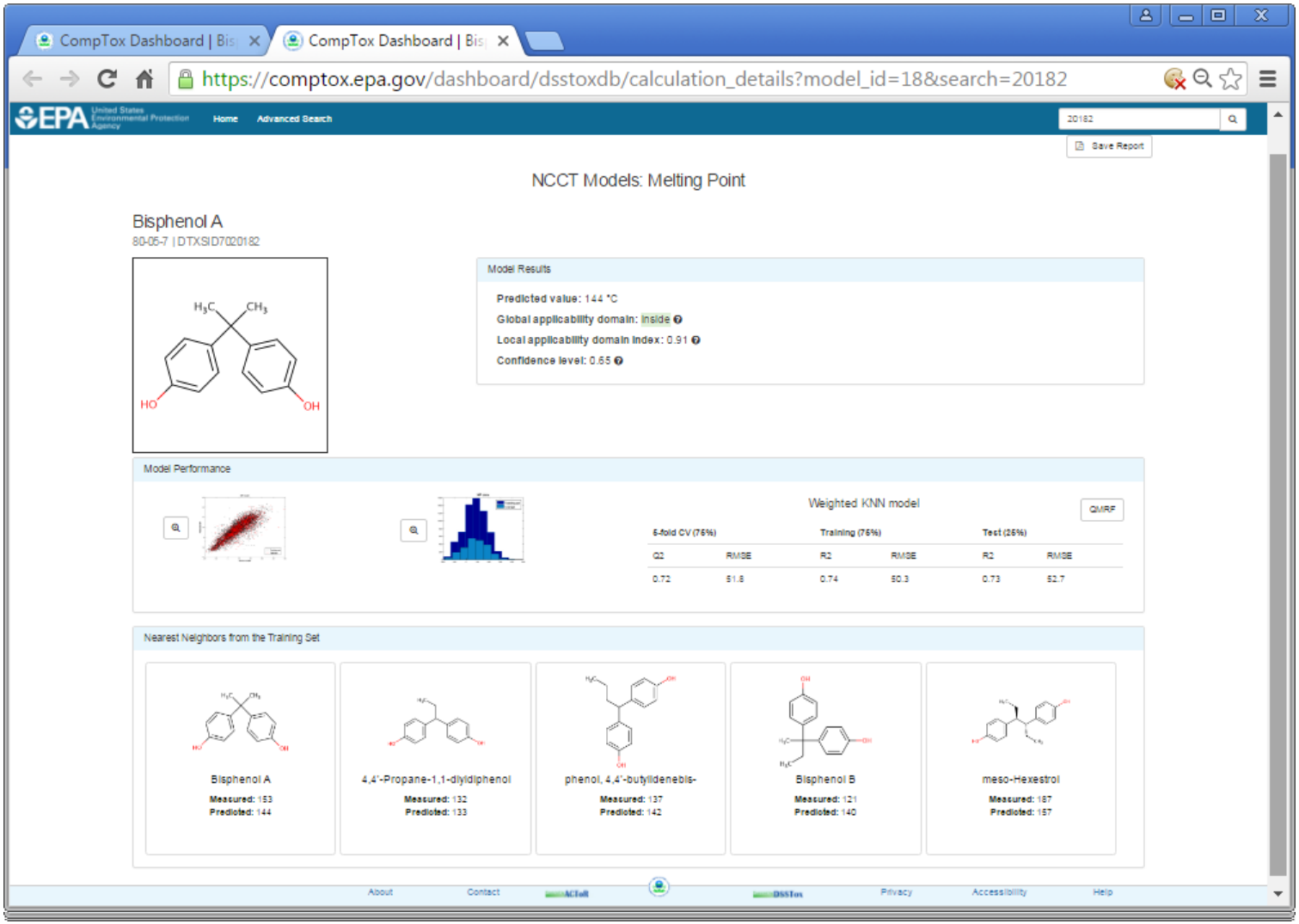
Extra slides

Comparing Dosimetry Adjusted Bioactivity with Exposure



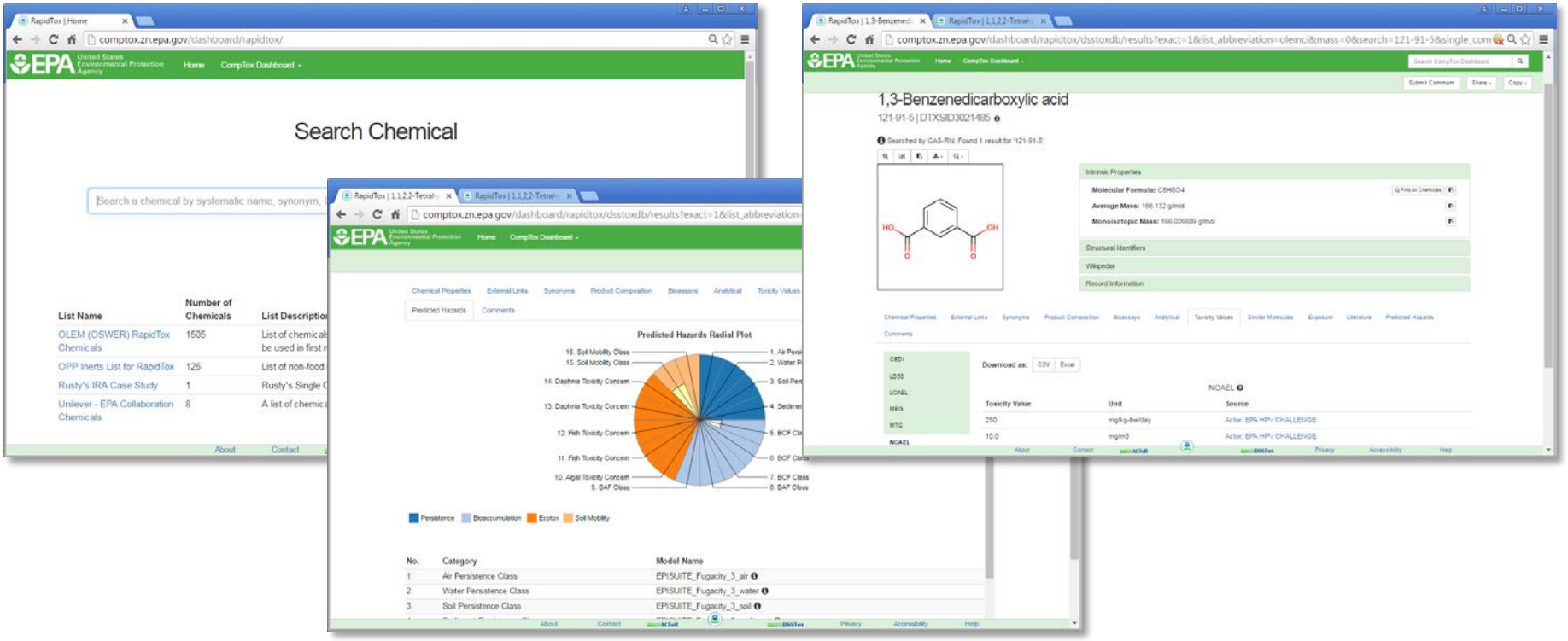
Wetmore *et al.*, *Tox Sci.*, 2012

New Chemistry Dashboard Delivers Structural and Property Data



https://comptox.epa.gov/dashboard

RapidTox Decision Support Dashboard in Development



- Semi-automated decision support tool with dashboard interface for high-throughput risk assessments
- Combining diverse data streams into quantitative toxicity values with associated uncertainty

"Do not let the perfect be the enemy of the good" *Voltaire*

"Do not let the perfect get in the way of developing and using in vitro data for use in risk assessments" *Crofton*