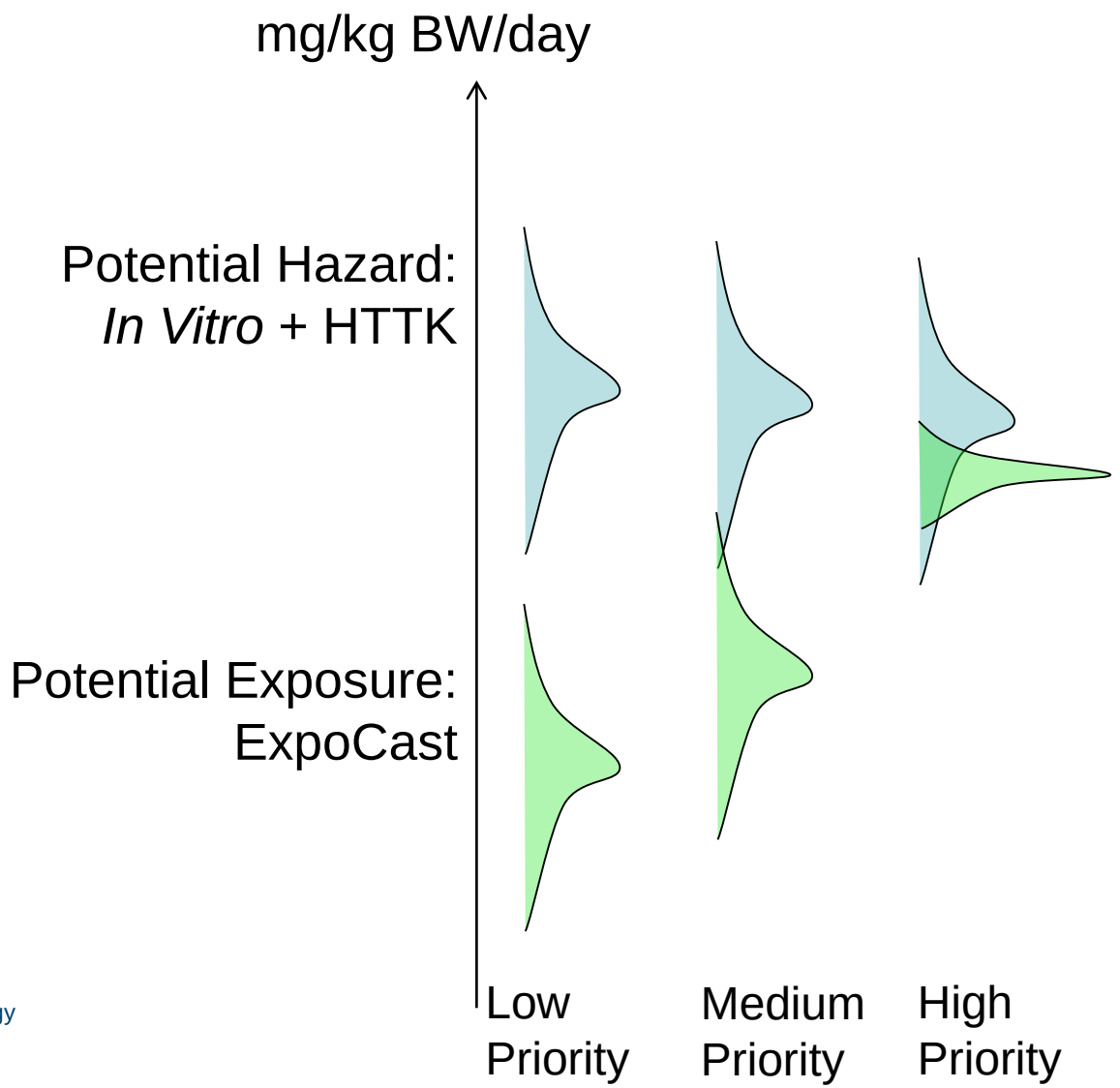
Using Tox21 in vitro data for hazard identification, development of prioritization-appropriate points of departure and chemical-class read-across applications

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

UCLA Alternatives Assessment Webinar Series 2015

- *In Vitro* assays: Bioactivity Concentration
- Need Bioactivity Dose to compare with exposure
- Convert using High Throughput Toxicokinetics (HTTK)

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



High-Throughput Risk Assessment (HTRA)

- Risk assessment approach
  - Estimate upper dose that is still protective
  - RfD, BMD are standard, animal-based quantities
  - Compare to estimated steady state exposure levels

- Contributions of high-throughput methods
  - Focus on molecular pathways whose perturbation can lead to adversity
  - Screen hundreds to thousands of chemicals in *in vitro* assays for those targets
  - Estimate oral dose using H-T pharmacokinetic modeling

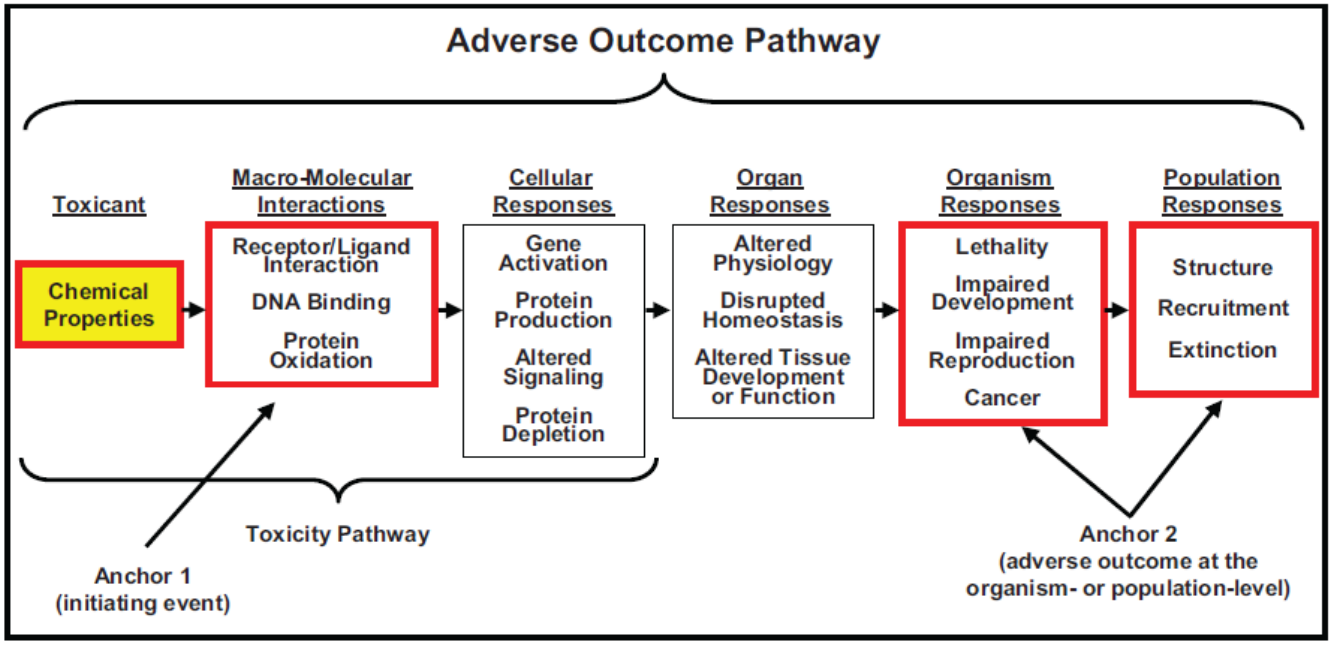
- Incorporate population variability and uncertainty

HTRA Basic Outline

1. Define molecular pathways linked to adverse outcomes
2. Measure activity *in vitro* in concentration-response (PD)
3. Estimate external dose to internal concentration scaling (PK)
4. Estimate dose at which pathway is perturbed *in vivo*
5. Estimate population variability and uncertainty in PK and PD
6. Estimate lower end of dose range for perturbation of pathway

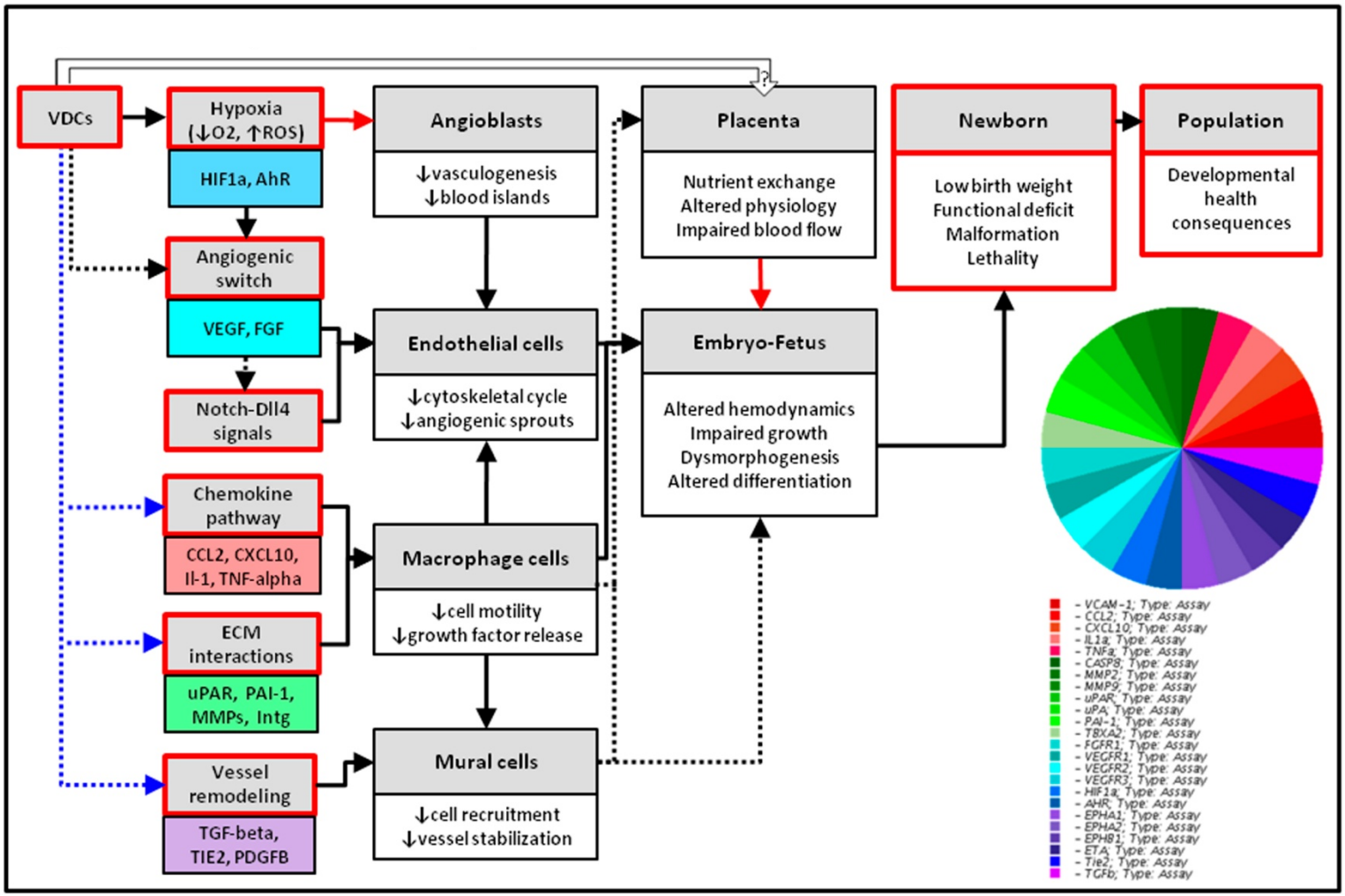
AOP / MOA Development

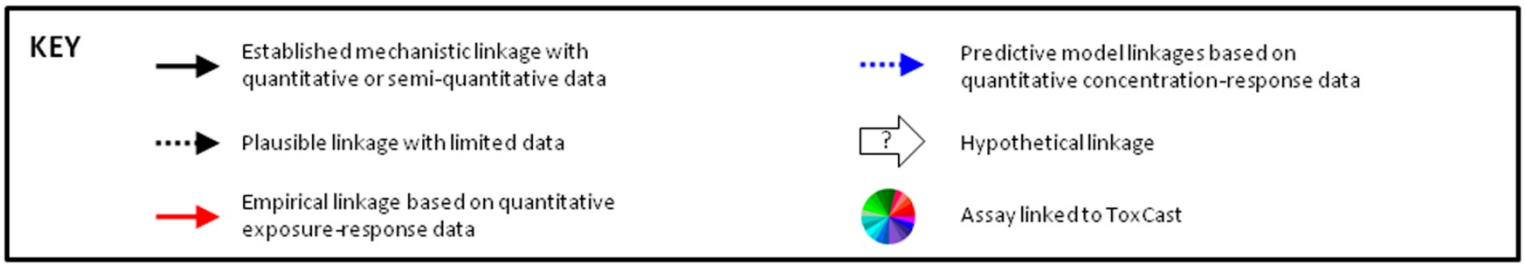
- International workgroups developing frameworks and models
  - OECD – AOP
  - WHO – MOA
- Key Concepts
  - Molecular Initiating Events or Key Events – measureable *in vitro*
  - Causal evidence for downstream effects
  - AOP includes effects up to the population level



Ankley et al. 2010

Example AOP: Embryonic Vascular Disruption





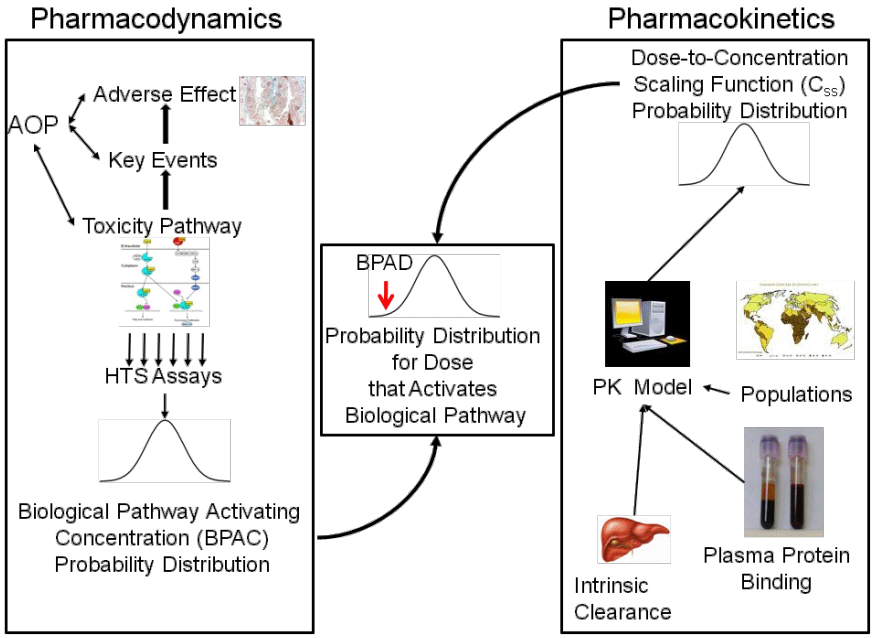
Knudsen and Kleinstreuer. Birth Def Res C. 2012

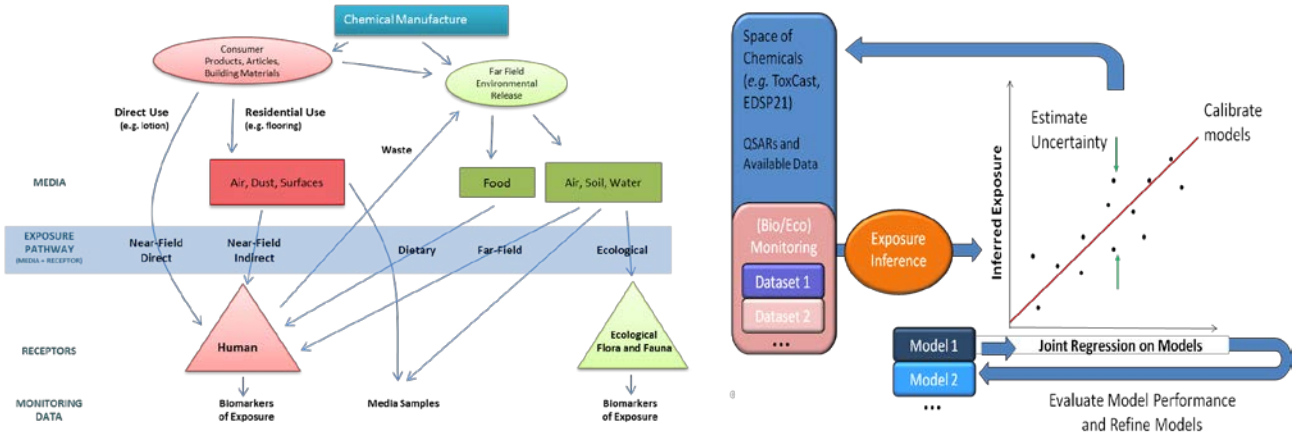
HTRA – High-Throughput Risk Assessment

High-throughput Hazard and Kinetics

+

High-throughput Exposure

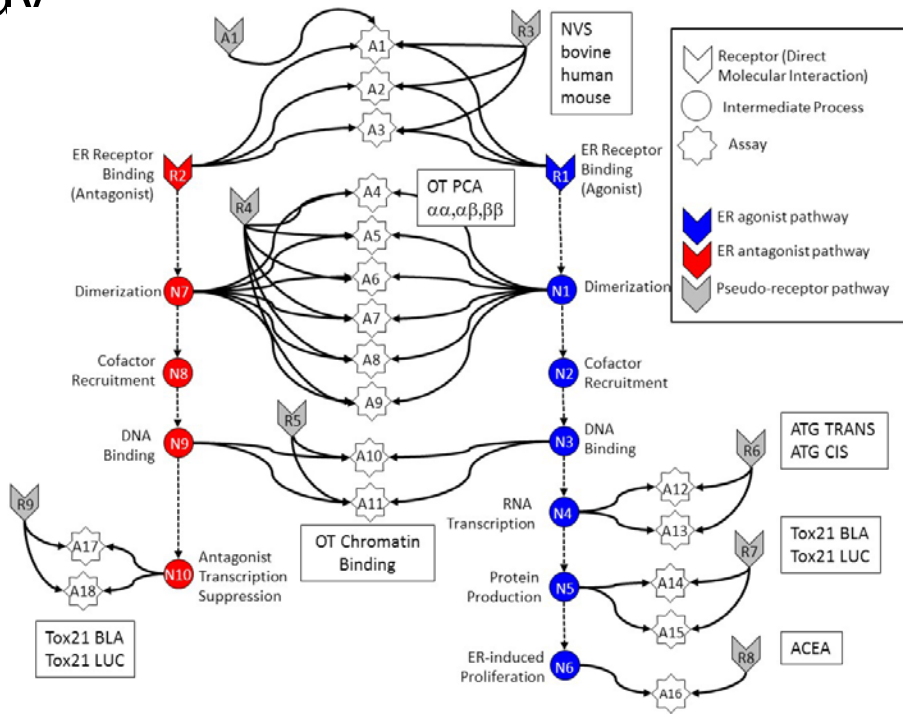




In Vitro Estrogen Receptor Model

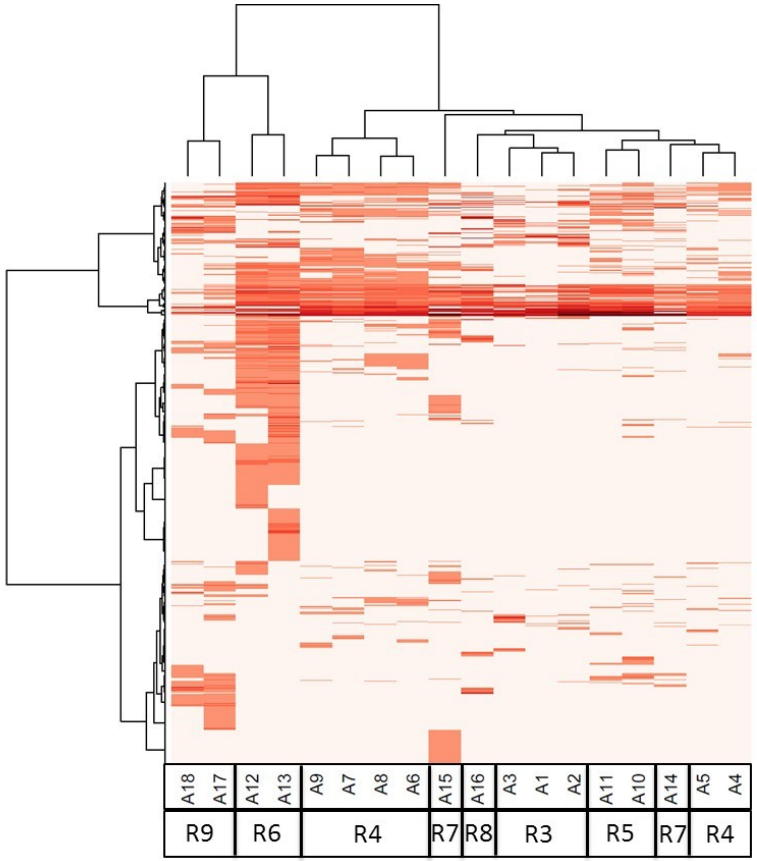
Combines results from multiple in vitro assays

- Use multiple assays per pathway
  - Different technologies
  - Different points in pathway
- No assay is perfect
  - Assay Interference
  - Noise
- Use model to integrate assays
- Evaluate model against reference chemicals
- Methodology being applied to other pathways



Major theme – all assays have false positives and negative

Assays cluster by technology, suggesting technology-specific non-ER bioactivity

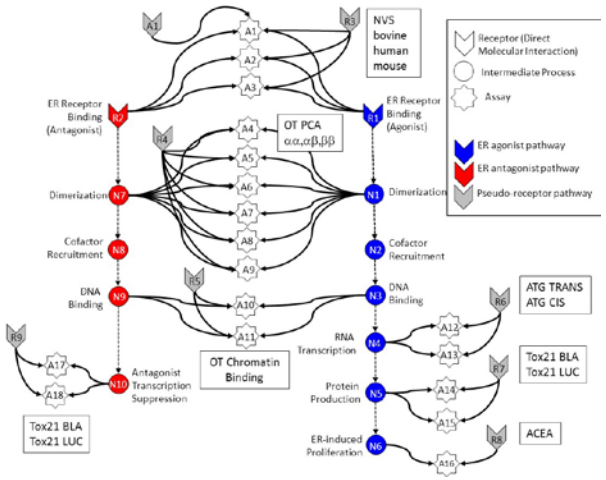


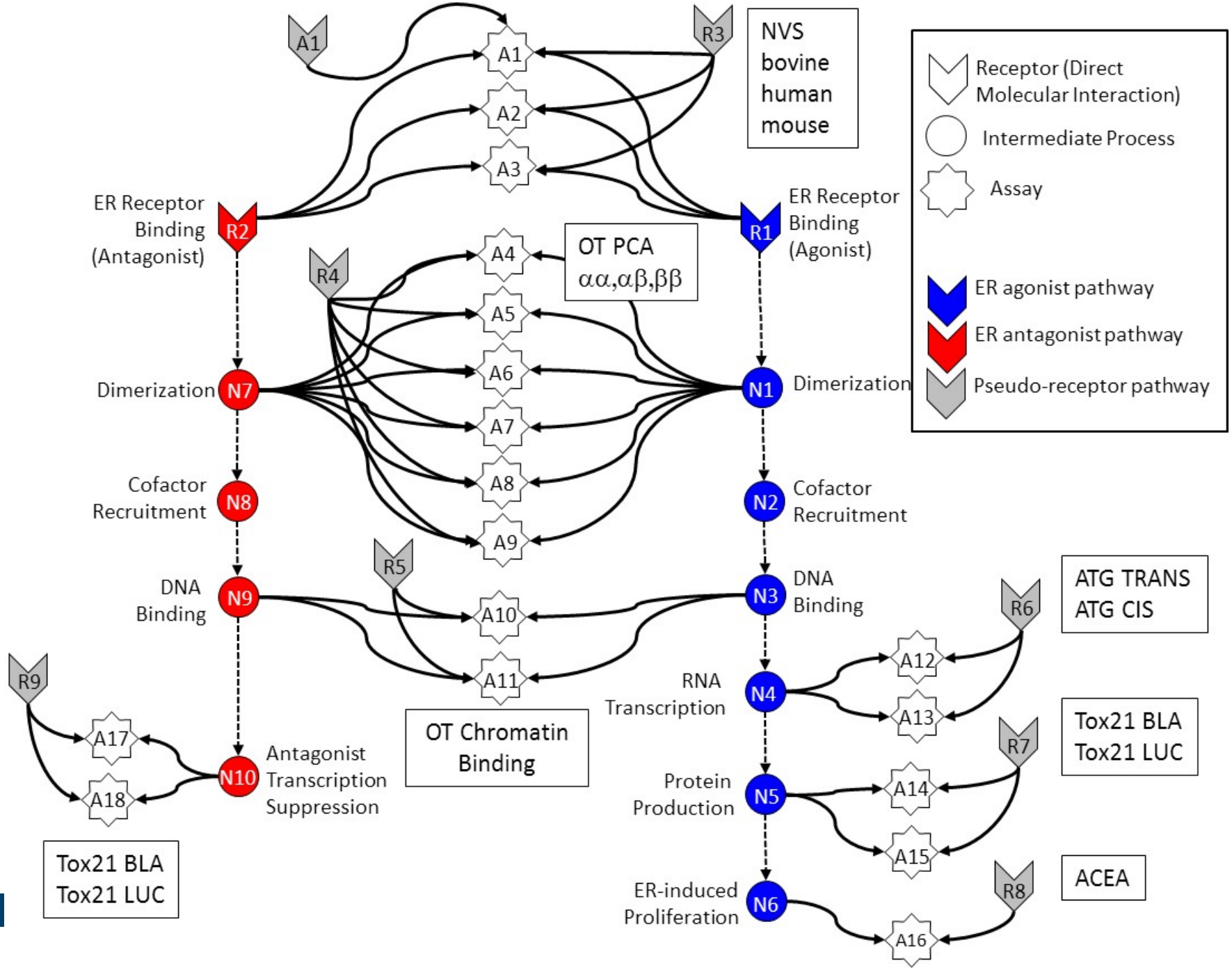
Much of this “noise” is reproducible

- “assay interference”
- Result of interaction of chemical with complex biology in the assay

EDSP chemical universe is structurally diverse

-Solvents
-Surfactants
-Intentionally cytotoxic compounds
-Metals
-Inorganics
-Pesticides
-Drugs




NVS
bovine
human
mouse
Receptor (Direct Molecular Interaction)
Intermediate Process
Assay
ER agonist pathway
ER antagonist pathway
Pseudo-receptor pathway
A1
A2
A3
R1
R2
R3
ER Receptor Binding (Antagonist)
ER Receptor Binding (Agonist)
OT PCA
αα,αβ,ββ
R4
A4
A5
A6
A7
A8
A9
Dimerization
N7
N1
Dimerization
Cofactor Recruitment
N8
N2
Cofactor Recruitment
R5
A10
A11
DNA Binding
N9
N3
DNA Binding
OT Chromatin Binding
ATG TRANS
ATG CIS
R6
A12
A13
RNA Transcription
N4
Tox21 BLA
Tox21 LUC
R7
A14
A15
Protein Production
N5
ACEA
R8
A16
ER-induced Proliferation
N6
R9
A17
A18
N10
Antagonist Transcription Suppression
Tox21 BLA
Tox21 LUC

Computational Model

$$A_i = \sum_j F_{ij} R_j$$

$\boldsymbol{A_i}$ is the efficacy of the assay at a given concentration
$\boldsymbol{R_j}$ is the "true" efficacy which is unobservable
$\boldsymbol{F}$ links receptors to assays

$$\varepsilon^2 = \sum_i (A_i^{pred} - A_i^{meas})^2 + penalty(\vec{R})$$

Solve a constrained least-squares problem to minimize difference between the measured and predicted assay values

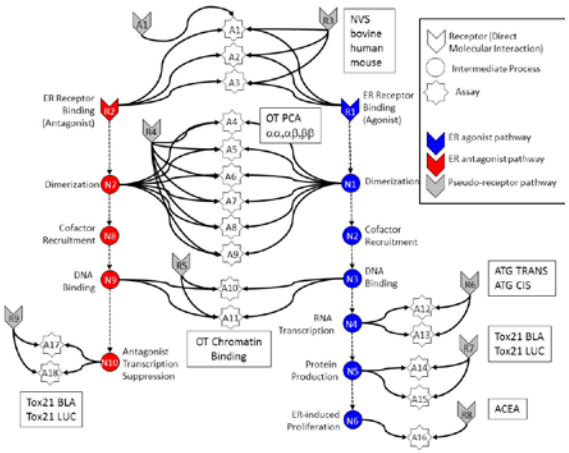
$$A_i^{pred} \in [1,0]$$

$$penalty(\vec{R}) = \alpha \frac{SR^2}{SR^2 + SR_0{}^2}$$

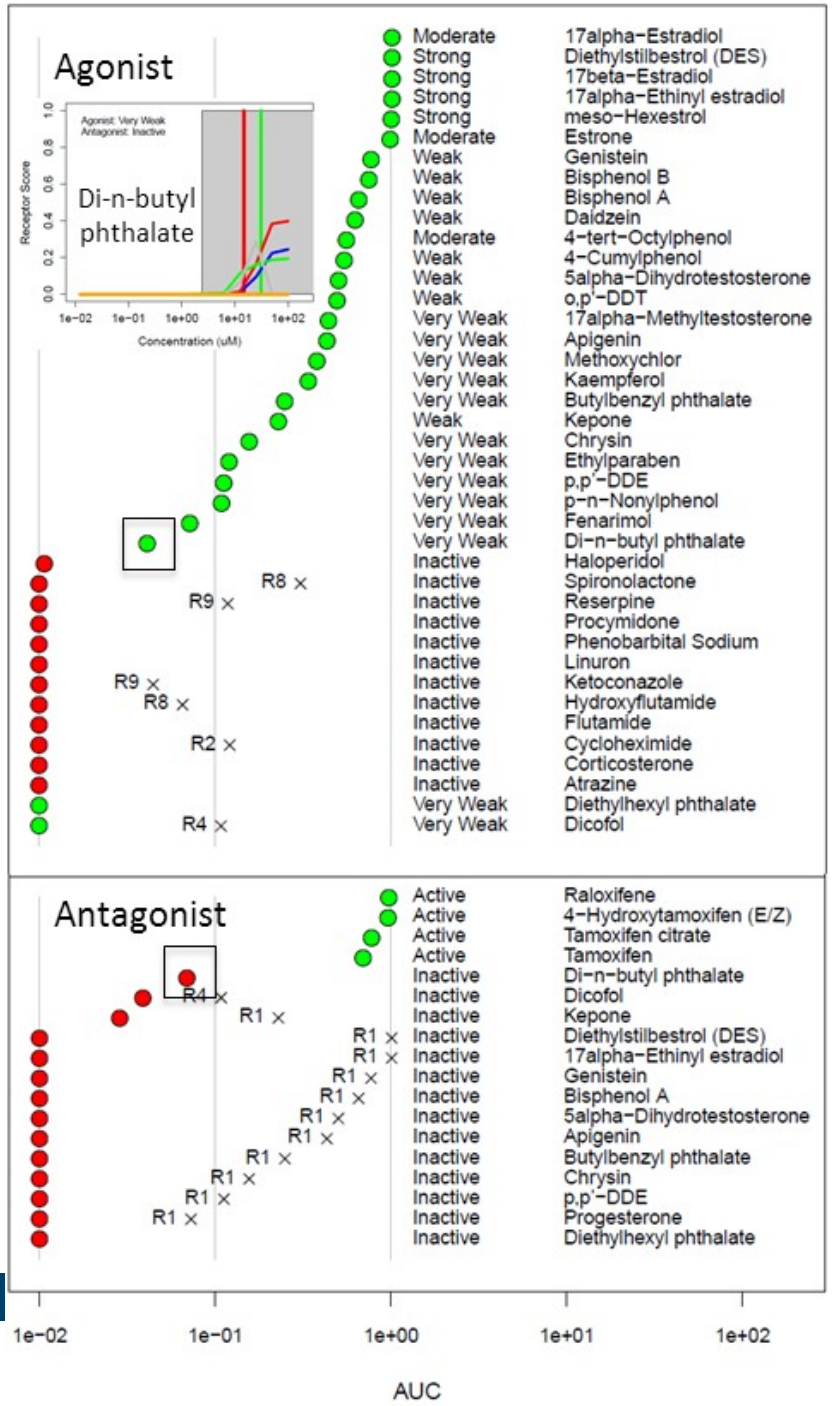
Penalty enforces physical assumption that chemical will not hit many targets simultaneously

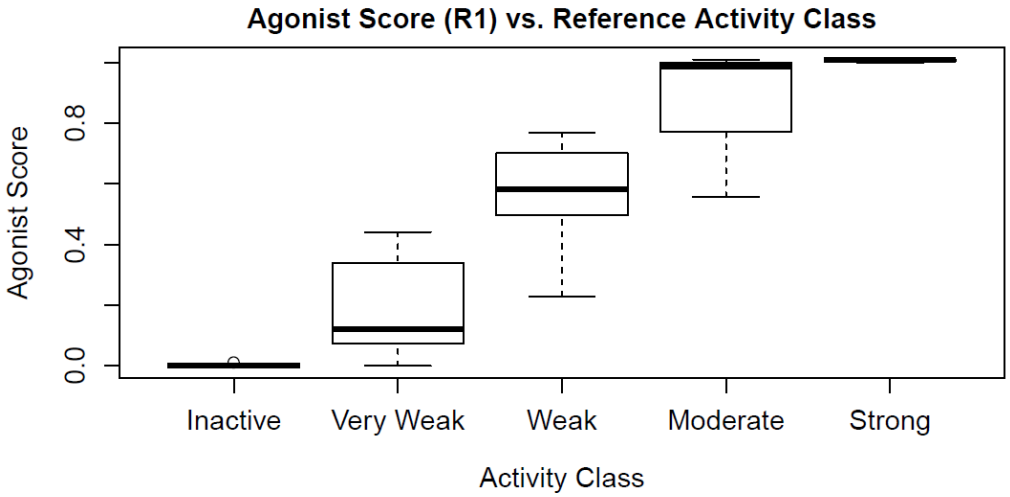
$$AUC_j = \frac{1}{N_{conc}} \sum_{i=1}^{N_{conc}} sign(slope) \times R_j(conc_i)$$

AUC Summarizes results



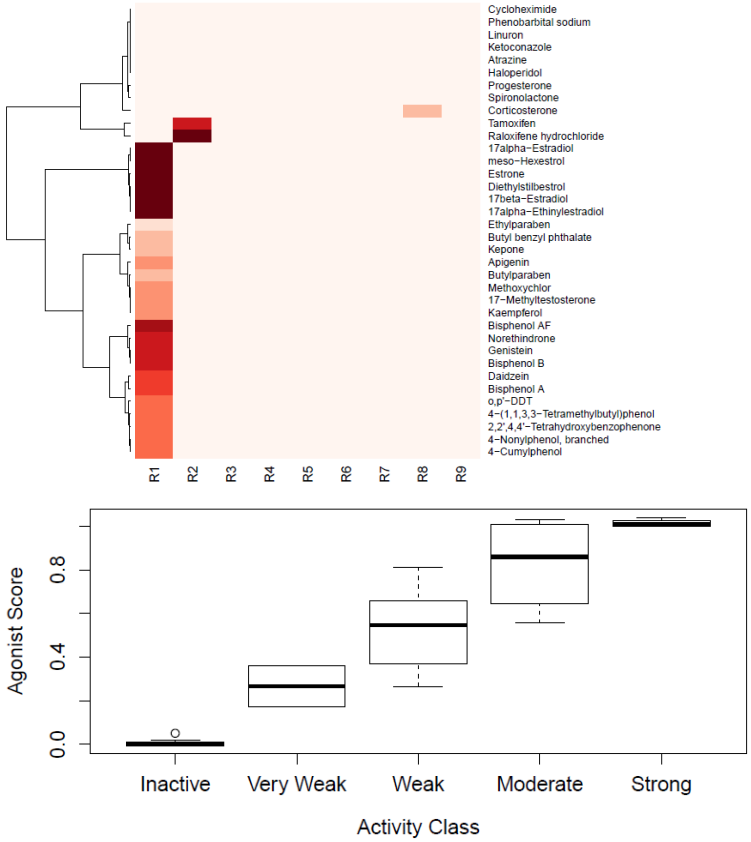
Reference Chemical Performance





ER Model Results

Appropriate Results for Reference Chemicals



Results For EDSP Universe Chemicals

1431 EDSP chemicals run *in vitro*
71 (5%) have a significant ER score

Mostly known chemical classes:

- Phenols
- Steroids
- Parabens
- Phthalates
- Organo-chlorides

Uses:

- Pesticides
- Pharmaceuticals
- Plastics
- Dyes
- Industrial Intermediates

CERAPP: Extend *In Vitro* data with QSAR Models

- Collaborative Estrogen Receptor Activity Prediction Project
- Goals:
  - Use ToxCast ER score (or other data) to build many QSAR models
  - Use consensus of models to prioritize chemicals for further testing
- Assumptions
  - ToxCast chemicals cover enough of chemical space to be a good "global" training set
  - Consensus of many models will be better than any one individually
- Process
  - Curate chemical structures
  - Curate literature data set
  - Build many models
  - Build consensus model
  - Evaluate models and consensus

Chemicals for Prediction: The Human Exposure Universe

- Estimate universe of man-made chemicals with potential for exposure
  - EDSP Universe (10K)
  - Chemicals with known use (40K)
    - From Chemical and Product Category DB (CPCat)
    - http://actor.epa.gov/cpcat
  - Canadian Domestic Substances List (DSL) (23K)
  - EPA DSSTox – structures of EPA/FDA interest (15K)
  - ToxCast and Tox21 (In vitro ER data) (8K)
- Unique set of structures: ~32K

Evaluation & Consensus

Models received:

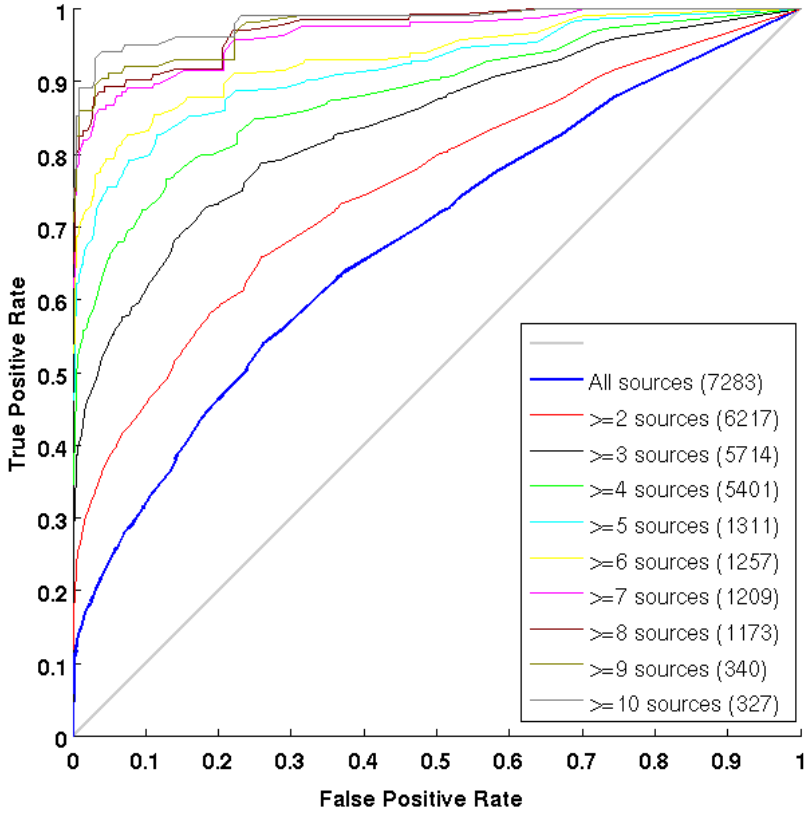
- Qualitative:
  - Binding: **22 models**
  - Agonist: **11 models**
  - Antagonist: **9 models**
- Quantitative:
  - Binding: **3 models**
  - Agonist: **3 models**
  - Antagonist: **2 models**

Evaluation procedure:

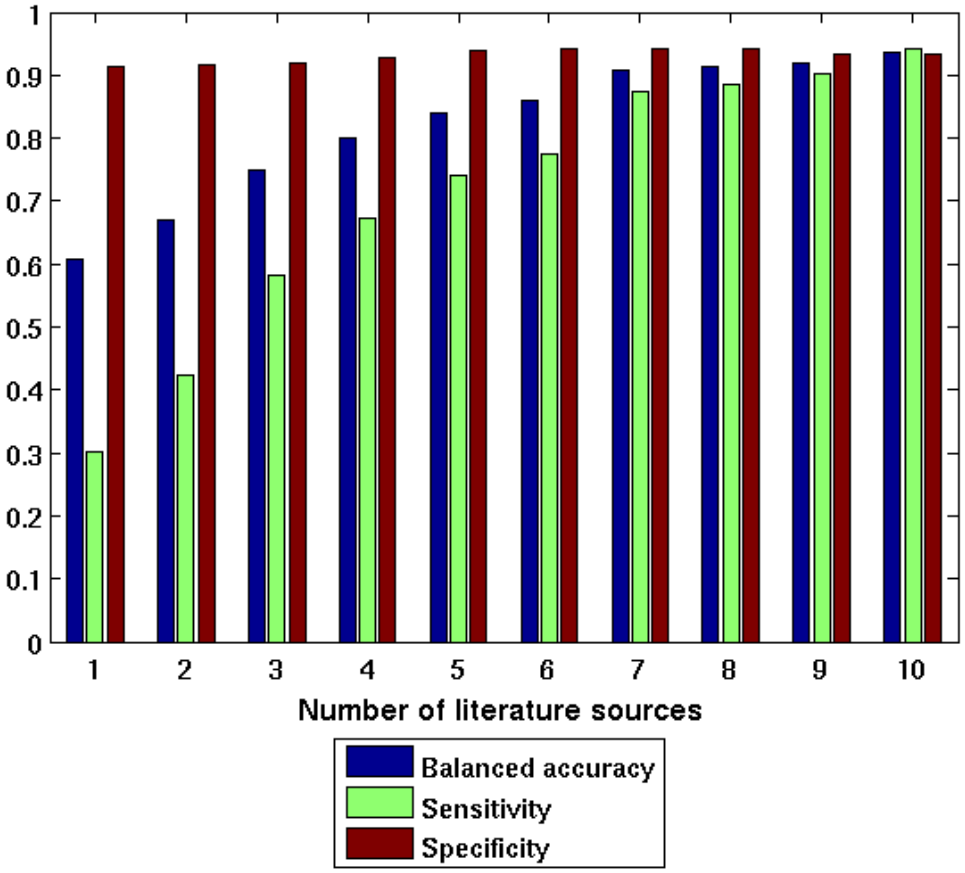
- On the EPA training set (1677)
- On the full evaluation set (~7k)
- Evaluation set with multi-sources
- Remove “Very Weak”
- Remove single source
- Remove chemicals outside the AD

Models provided by 17 groups in U.S. and Europe

Consensus evaluation



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



Total Database
Binders: 3961
Agonists: 2494
Antagonists: 2793

ER QSAR Summary

- Many ER QSAR and docking models built using ER model result (AUC) as training data
- 5-10% of chemical universe has predicted potential for ER bioactivity
- QSAR-model positives are candidates for follow-up in vitro testing

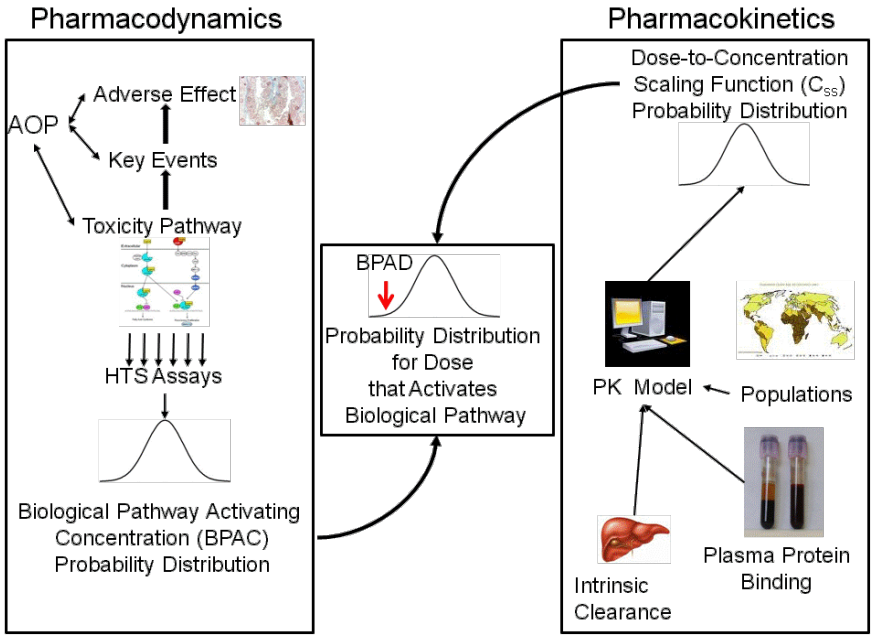
- Consensus of models gives high balanced accuracy for literature data that is internally consistent
- Open-source structure preparation process performed on all EDPS universe distinct chemicals (and others)

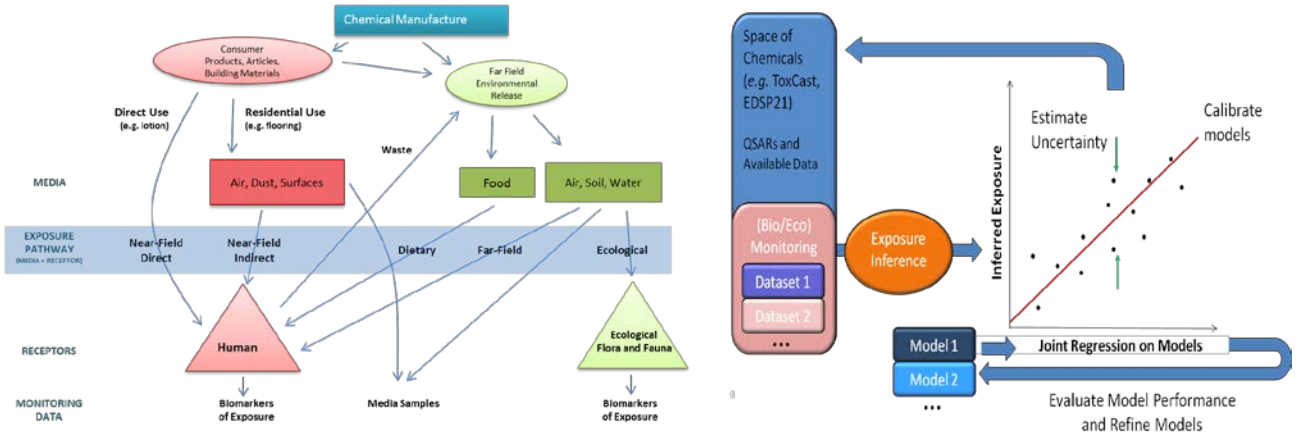
HTRA – High-Throughput Risk Assessment

High-throughput Hazard and Kinetics

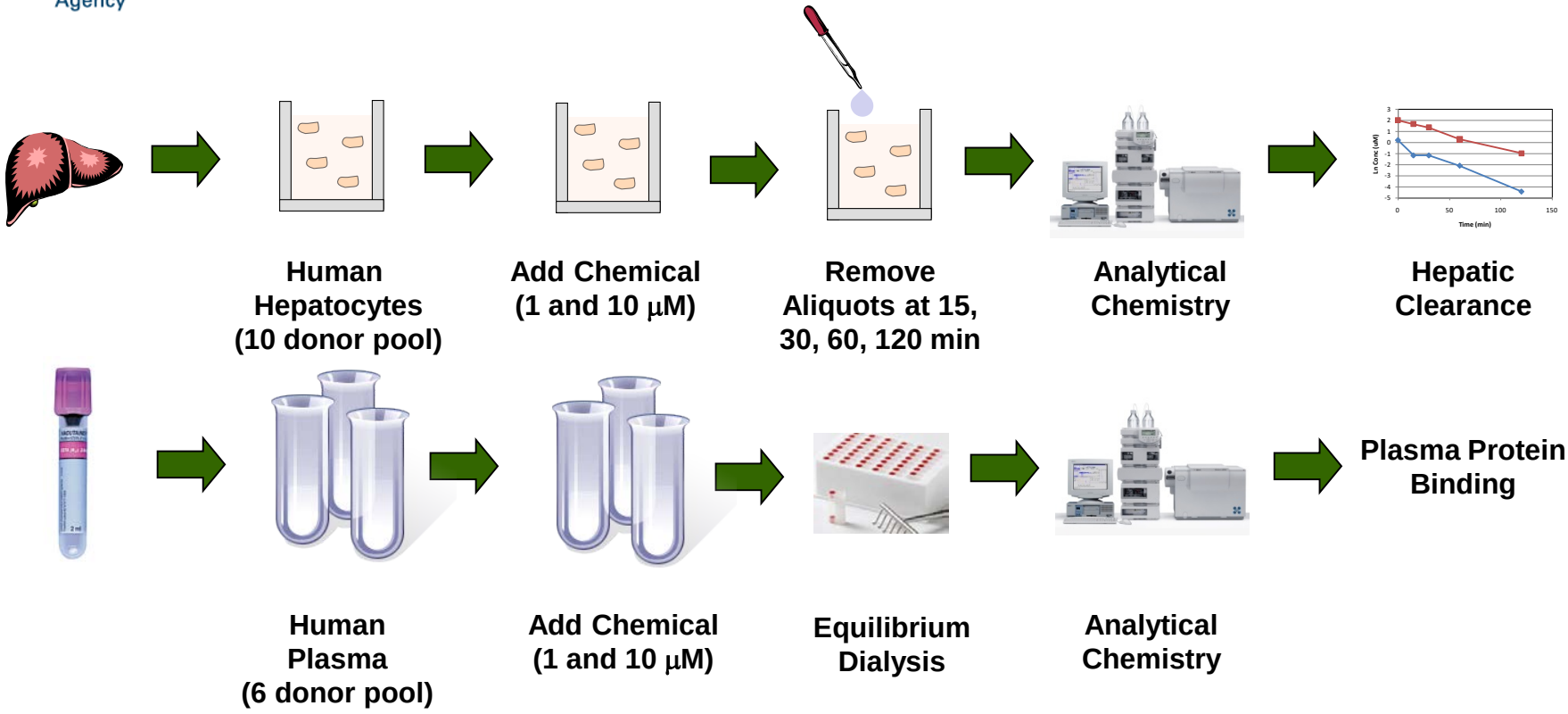
+

High-throughput Exposure





Adding Pharmacokinetics: High-Throughput ToxicoKinetics (HTTK)

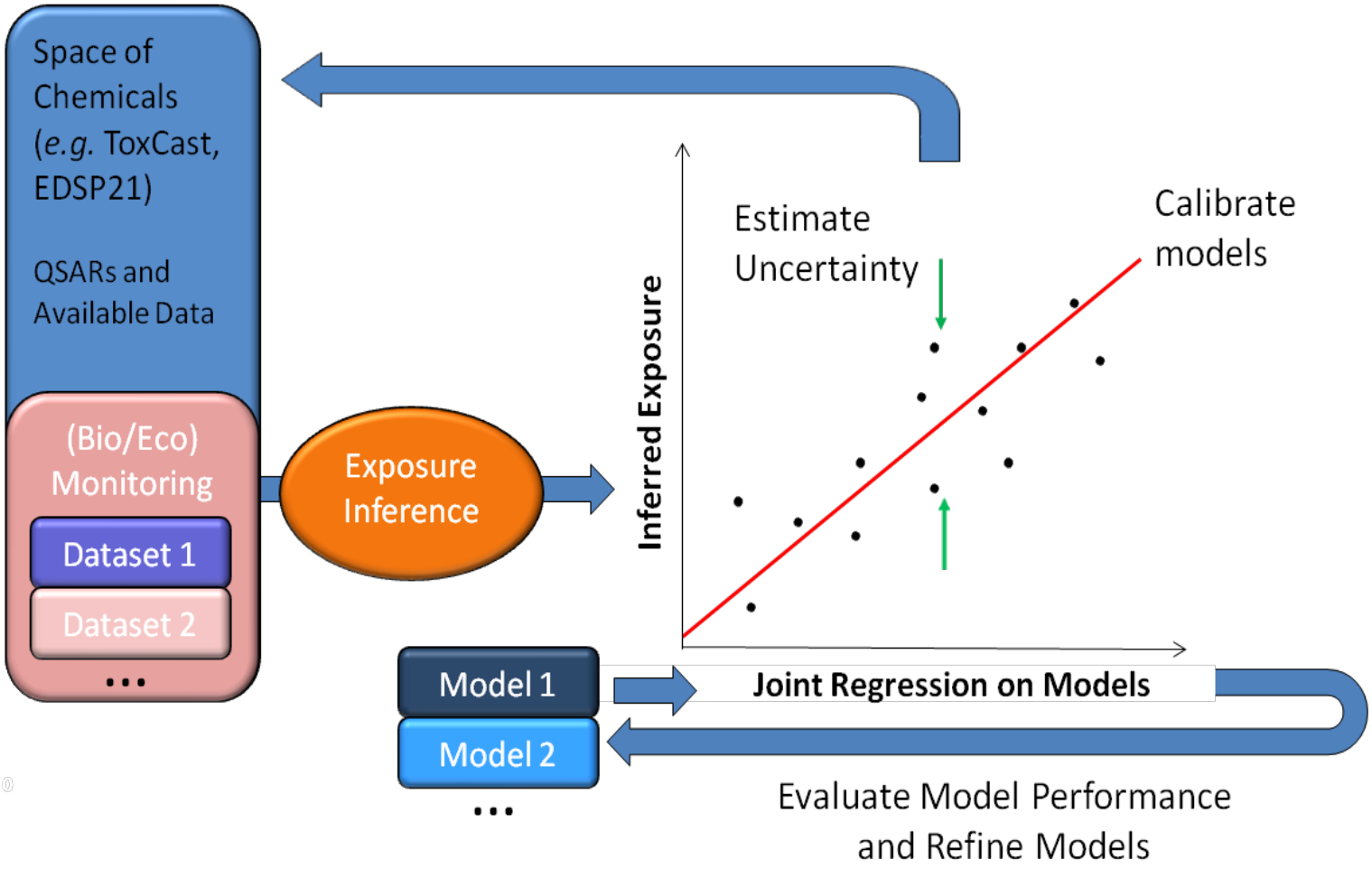


Combine experimental data w/ PK Model to estimate dose / concentration scaling

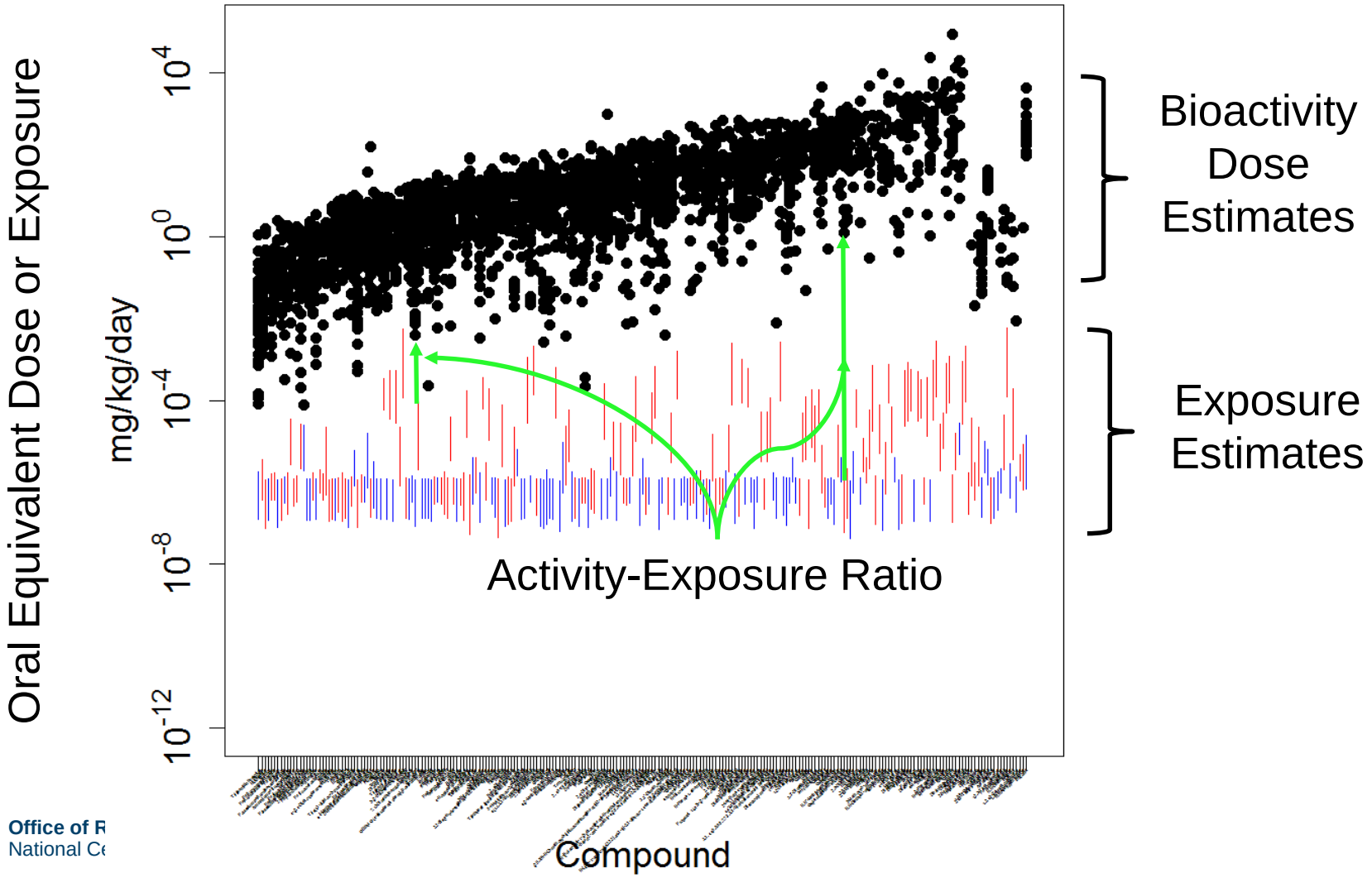
Bioactivity Dose = Bioactivity Concentration / Css

ExpoCast Exposure Modeling

Output: Estimate of exposure (w/ confidence interval)



Combine Hazard, HTTK Dose, Exposure
Output: "Activity Exposure Ratio (AER)"



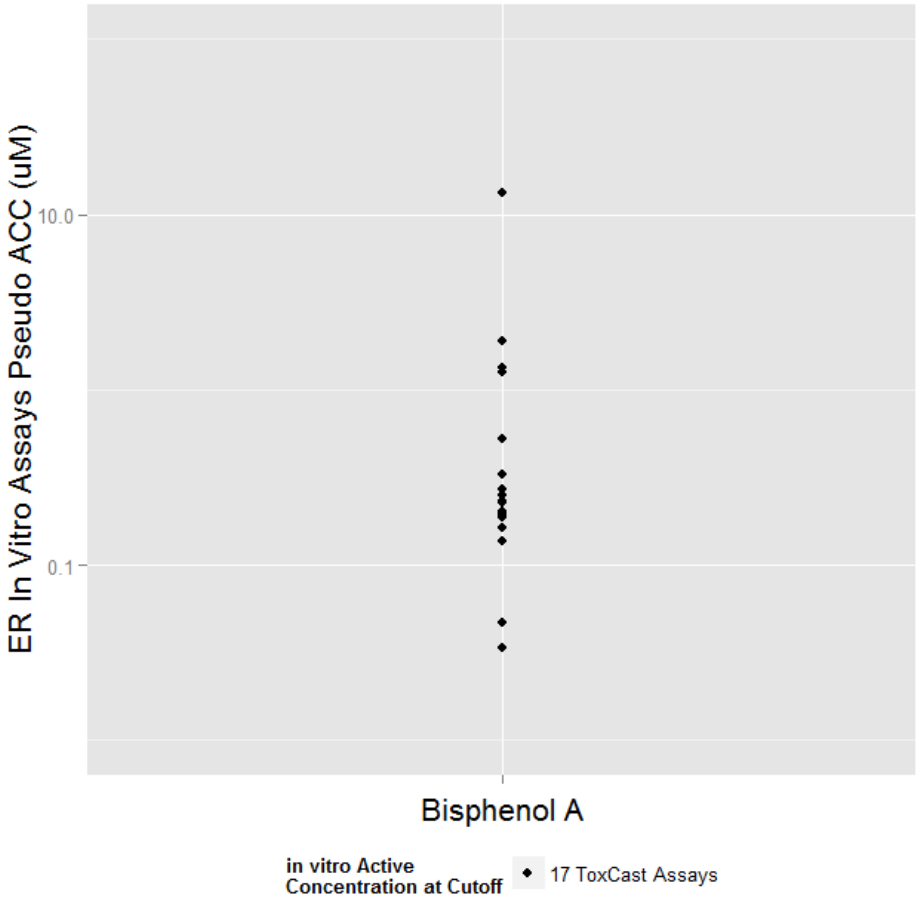
Summary of Uncertainty and Variability Components for HTRA

	Uncertainty	**Variability**
Pharmacodynamics Note: Data is human-derived	Data uncertainty (potency) Other biology not included	*Default for now* HapMap cell-line experiments may help
Pharmacokinetics Note: Data is human-derived	Data uncertainty (plasma protein binding, intrinsic clearance).	Model variability in liver function as *f*(age, sex, body weight)
Exposure Note: Model is parameterized using NHANES data	Includes uncertainty in biomonitoring data	NHANES-derived variability SHEDS-like models can be used

Summary: How Well Do We Understand Uncertainty and Variability?

	Uncertainty	Variability
Pharmacodynamics Note: Data is human-derived	Data uncertainty (potency) Other biology not included	*Default for now* HapMap cell-line experiments may help
Pharmacokinetics Note: Data is human-derived	Data uncertainty (plasma protein binding, intrinsic clearance).	Model variability in liver function as *f*(age, sex, body weight)
Exposure Note: Model is parameterized using NHANES data	Includes uncertainty in biomonitoring data	NHANES-derived variability SHEDS-like models can be used

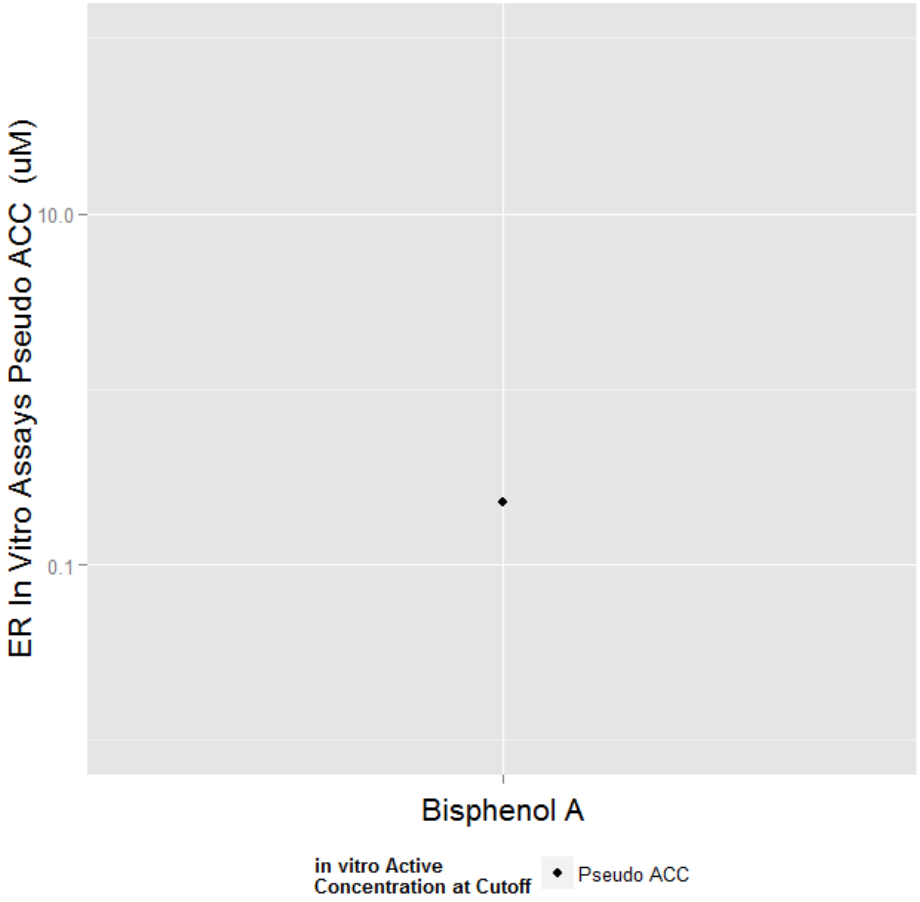
ER Case Study / BPA



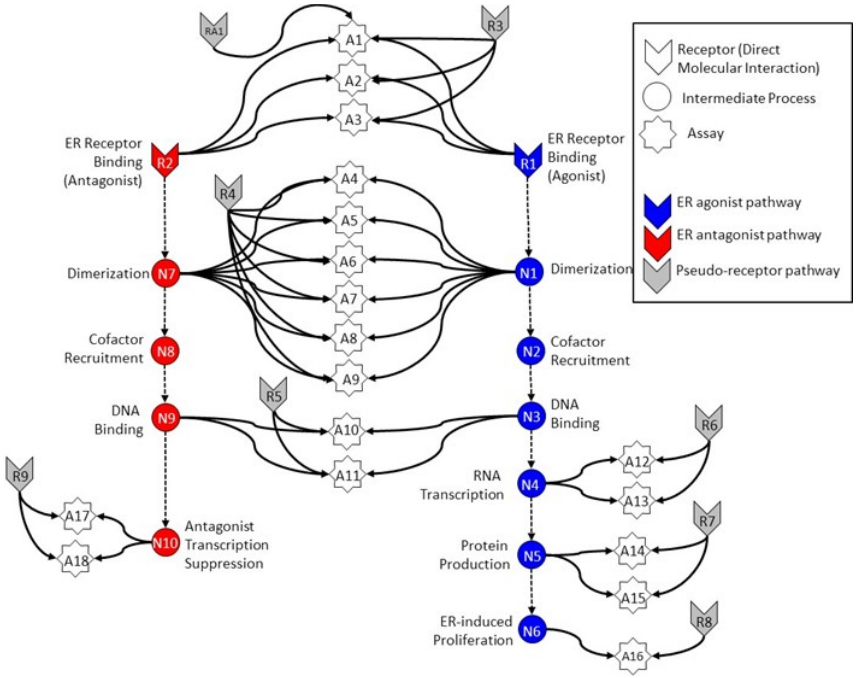
- Bisphenol A was active at some concentration for 17 of 18 ER-related assays

Assay	Conc.
NVS_NR_bER_ACC	0.19
NVS_NR_hER_ACC	0.20
NVS_NR_mERa_ACC	0.27
OT_ER_ERaERa_0480_ACC	1.27
OT_ER_ERaERa_1440_ACC	1.34
OT_ER_ERaERb_0480_ACC	0.23
OT_ER_ERaERb_1440_ACC	0.25
OT_ER_ERbERb_0480_ACC	0.23
OT_ER_ERbERb_1440_ACC	0.19
OT_ERa_EREGFP_0120_ACC	0.33
OT_ERa_EREGFP_0480_ACC	0.52
ATG_ERa_TRANS_up_ACC	0.03
ATG_ERE_CIS_up_ACC	0.05
Tox21_ERa_BLA_Agonist_ratio_ACC	1.88
Tox21_ERa_LUC_BG1_Agonist_ACC	0.14
ACEA_T47D_80hr_Positive_ACC	0.16
Tox21_ERa_BLA_Antagonist_ratio_ACC	13.27
Tox21_ERa_LUC_BG1_Antagonist_ACC	1000000

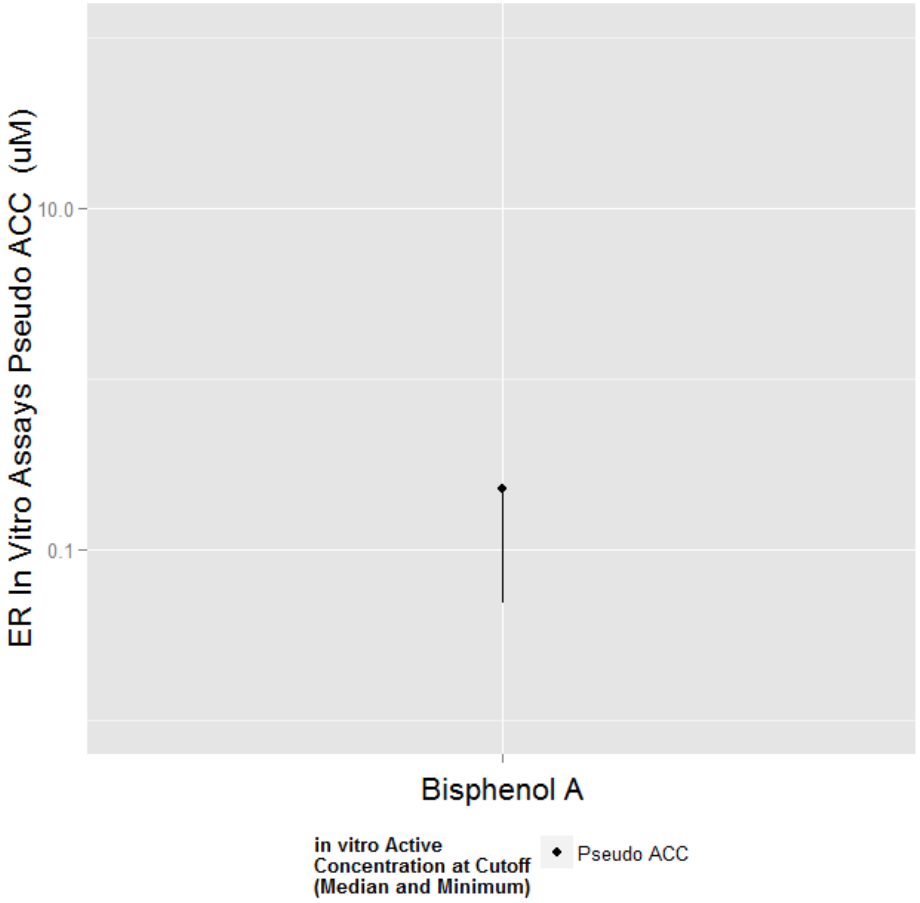
ER Case Study / BPA



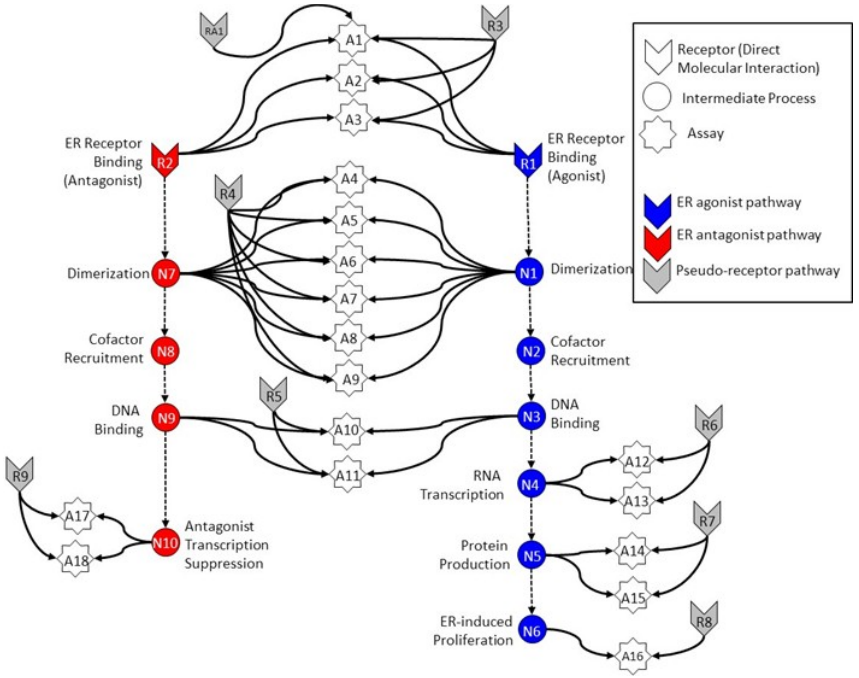
- A mathematical model was used to integrate all assays into a single predicted active concentration



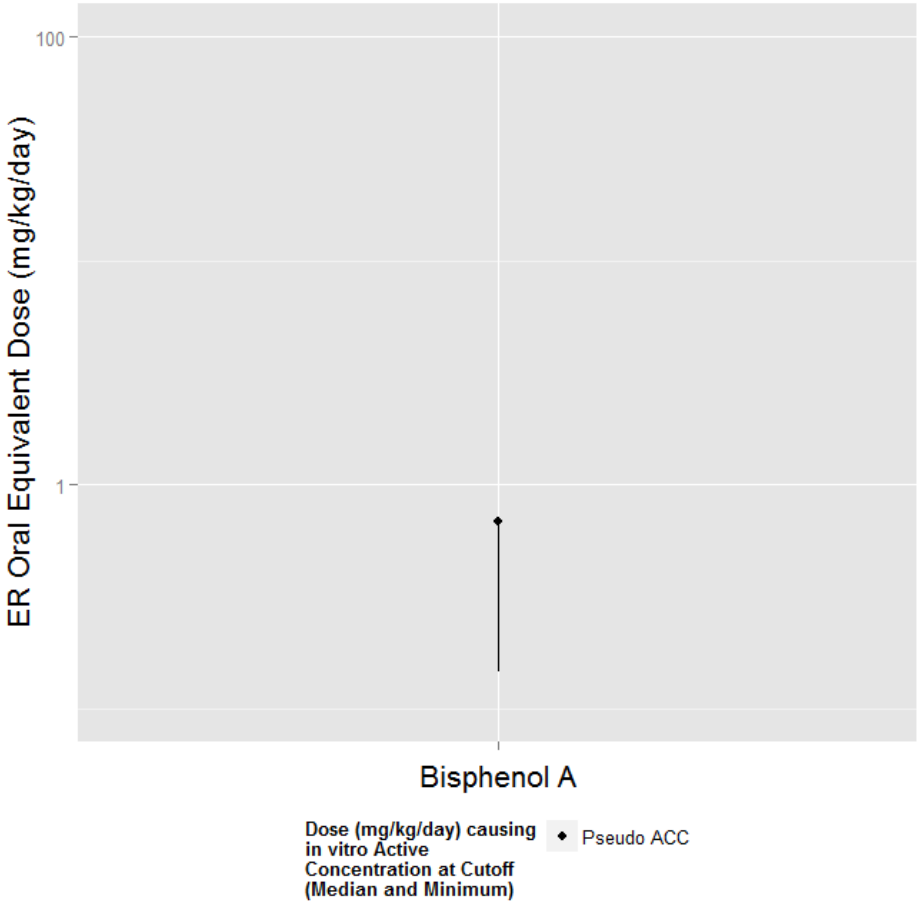
ER Case Study / BPA



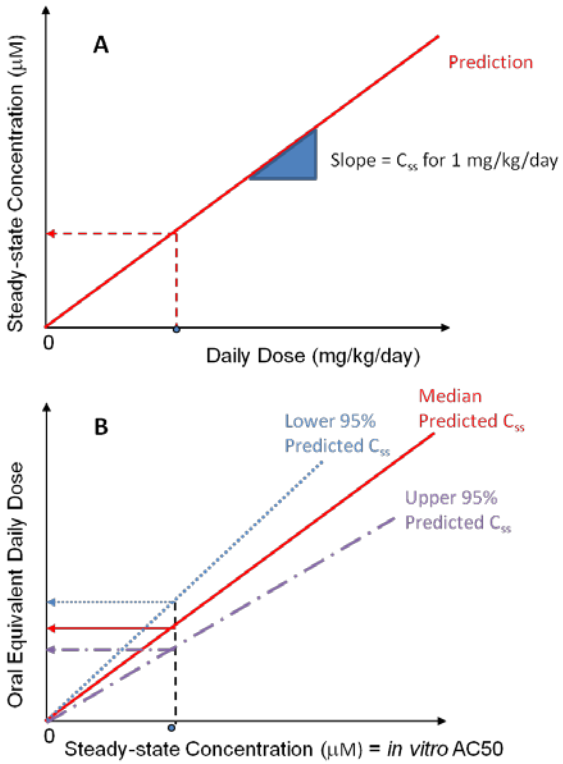
- The error bar indicates the span between the median and the minimum plausible active concentration



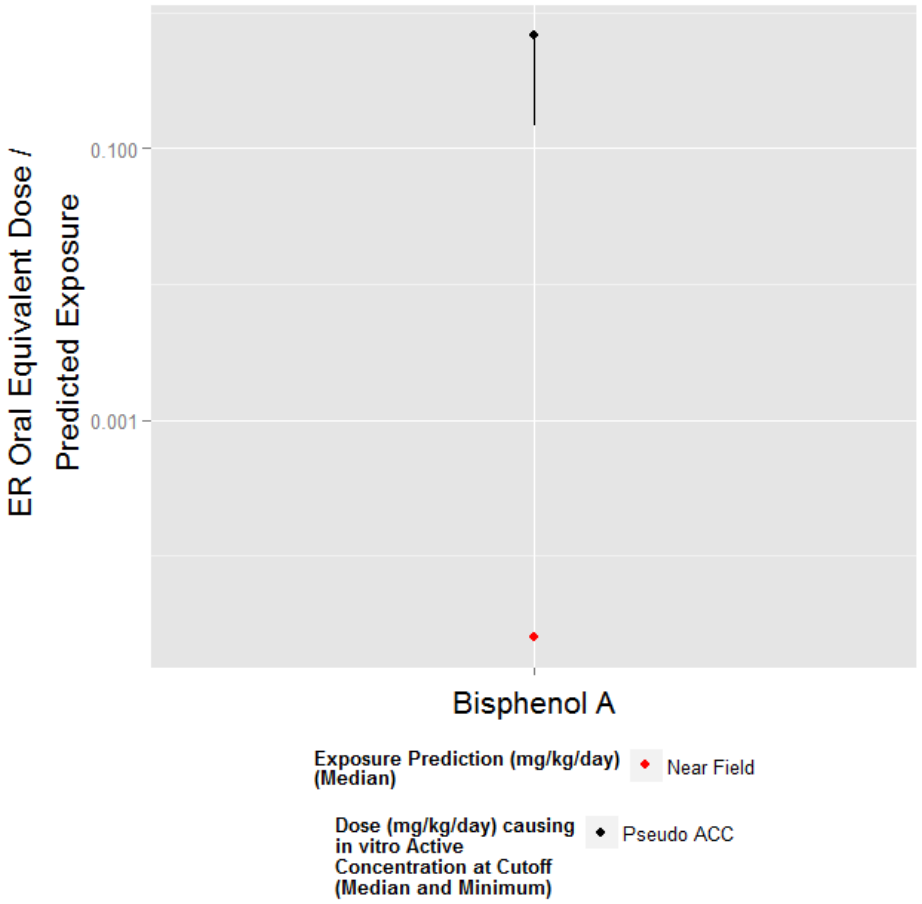
ER Case Study / BPA



- Reverse dosimetry based on HTTK data was used to predict an oral equivalent dose that would cause the ACC in plasma for the 95-percentile, most sensitive adult



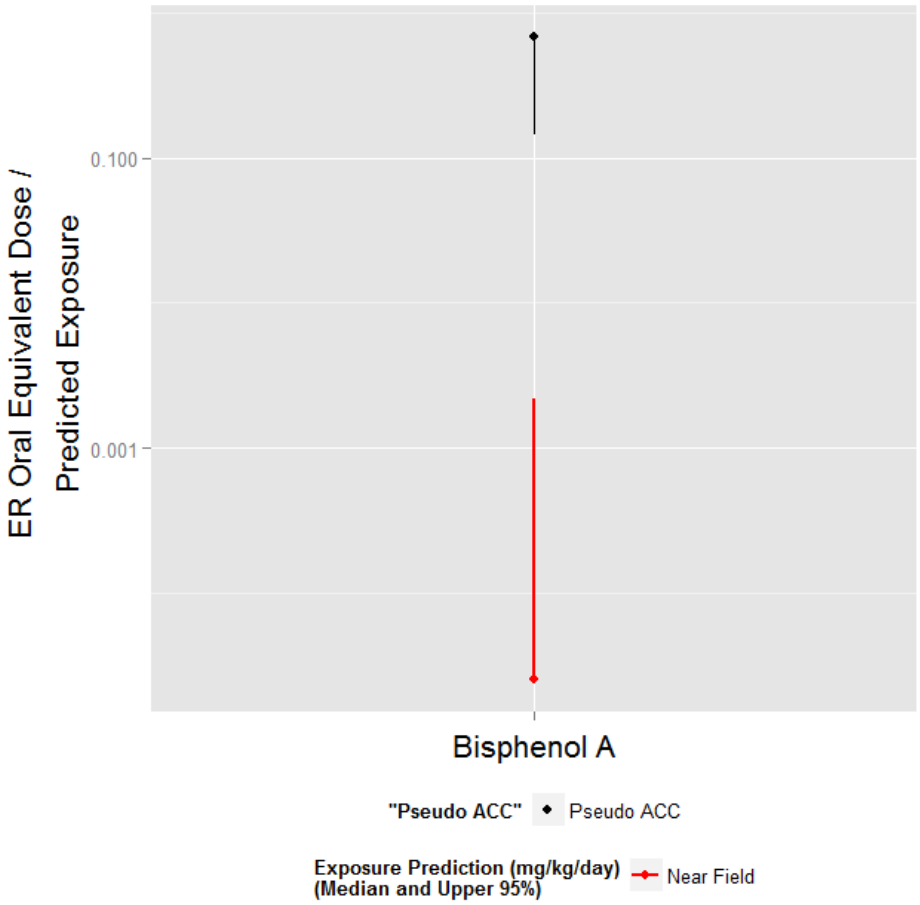
ER Case Study / BPA



- Based on the ACToR UseDB descriptors and production volume, a median exposure for similar NHANES chemicals can be predicted

Heuristic	Bisphenol A
Consumer & Industrial Use	Yes
Industrial Use Only	No
Pesticide Inert	No
Pesticide Active	No
Production Volume	> 1 billion lbs/year

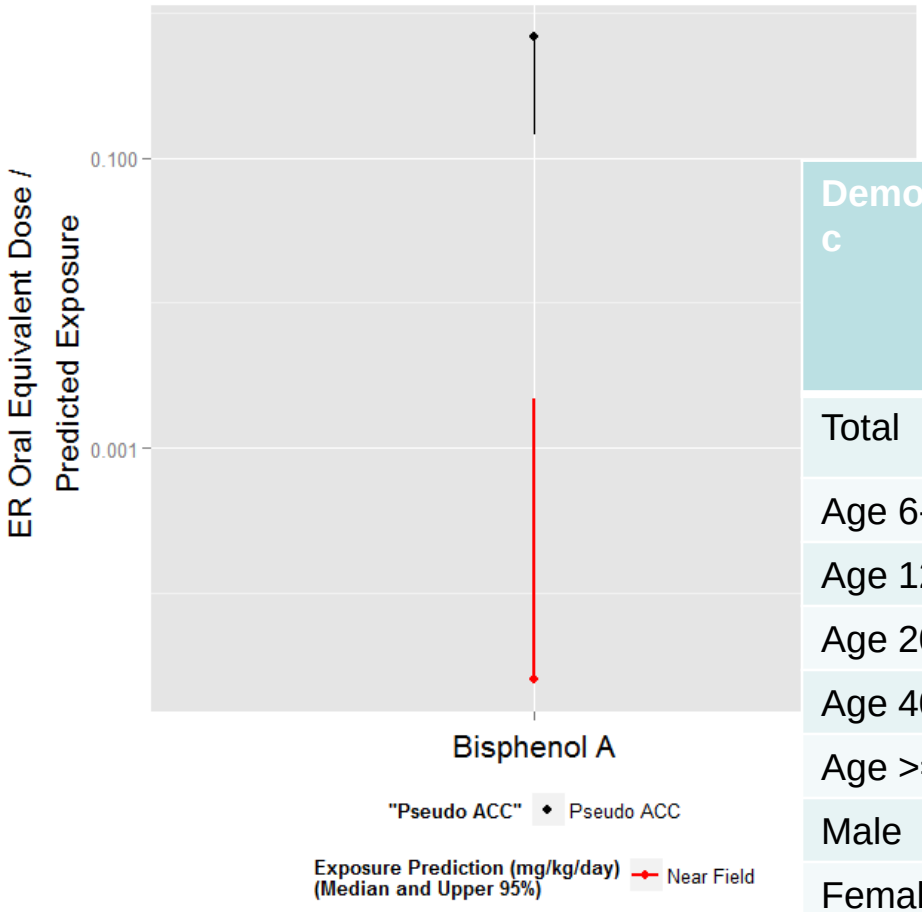
ER Case Study / BPA



- Due to the large uncertainty, the upper 95% limit of the exposure estimate credible interval is used

Heuristic	Bisphenol A
Consumer & Industrial Use	Yes
Industrial Use Only	No
Pesticide Inert	No
Pesticide Active	No
Production Volume	> 1 billion lbs/year

ER Case Study / BPA

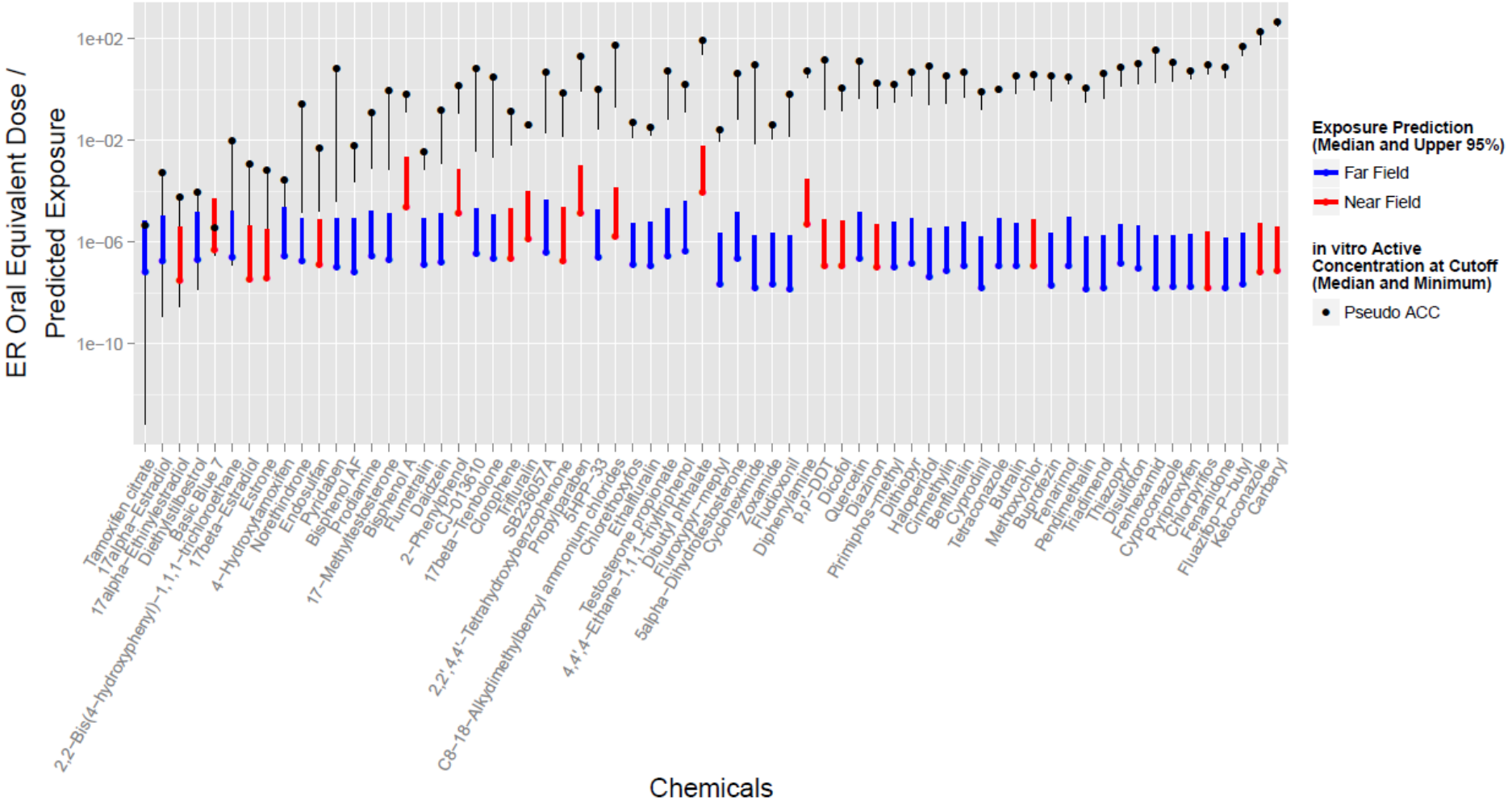


- LaKind and Naiman (2011) Estimated Exposure to BPA from NHANES data in ng/kgBW/day):

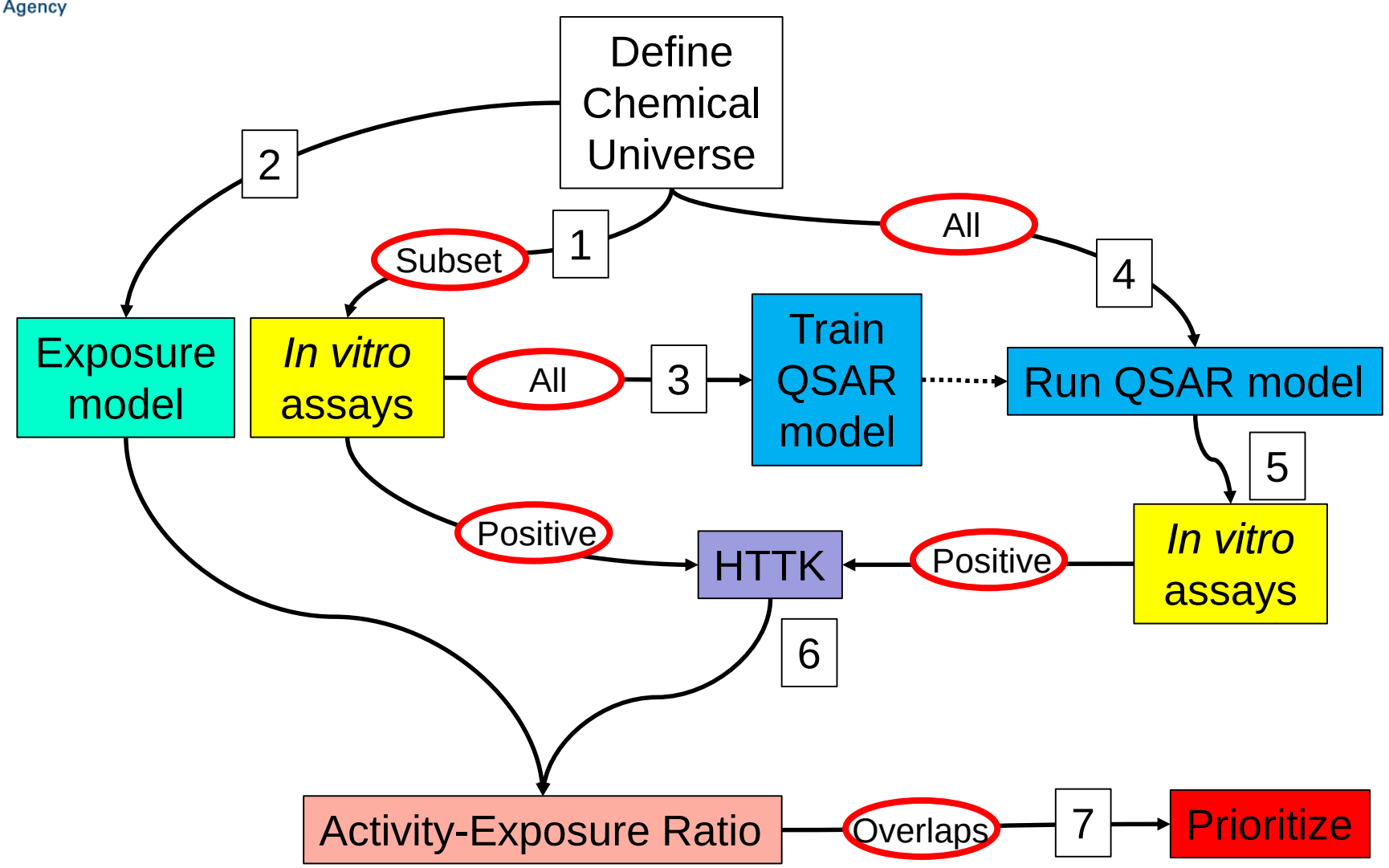
Demographic	LaKind and Naiman (2011)	ExpoCast Geometric Mean Median	ExpoCast Geometric Mean Upper 95%
Total	35.1	25.0	2193
Age 6-11y	54	63	4984
Age 12-19y	48	59	5169
Age 20-39y*	38.5	57	6056
Age 40-59y*	28.9	57	6056
Age >=60y	27.3	66	84221
Male	39.6	38	3132
Female	31.2	12	1125

*ExpoCast makes single prediction for Age 20-59y

HTRA for ER
Only show subset (of 290) with ER bioactivity



Summary: Overall Prioritization Scheme



Summary for ER HTRA Case Study

- Method ranks chemicals by their exposure-bioactivity dose differences
- Use *in vitro* assay data to derive a concentration at which pathway-based bioactivity occurs
- Use *in vitro* toxicokinetics to convert to an oral equivalent dose
- Use exposure models to estimate exposure, given assumptions about near-field use
- Conservative assumptions are used
- All quantities include estimates of uncertainty and variability
- Most chemicals showing overlap between exposure and likely bioactivity doses are drugs or natural hormones

Understanding Success and Failure

- Why *In vitro* to *in vivo* can work:
  - Chemicals cause effects through direct molecular interactions that we can measure with *in vitro* assays

- Why *in vitro* to *in vivo* does not always work:
  - Pharmacokinetics issues: biotransformation, clearance (**FP**, **FN**)
  - Assay coverage: don't have all the right assays (**FN**)
  - Tissue issues: may need multi-cellular networks and physiological signaling (**FN**)
  - Statistical power issues: need enough chemicals acting through a given MOA to be able to build and test model (**FN**)
  - Homeostasis: A multi-cellular system may adapt to initial insult (**FP**)
  - *In vitro* assays are imperfect (**FP**, **FN**)
  - *In vivo* rodent data is imperfect (**FP**, **FN**)

Acknowledgements

EPA NCCT
Rusty Thomas
Kevin Crofton
Keith Houck
Ann Richard
Richard Judson
Tom Knudsen
Matt Martin
Woody Setzer
John Wambaugh
Monica Linnenbrink
Jim Rabinowitz
Steve Little

EPA NCCT
Nancy Baker
Jeff Edwards
Dayne Filer
Jayaram Kancherla
Parth Kothiya
Jimmy Phuong
Jessica Liu
Doris Smith
Jamey Vail
Hao Truong
Sean Watford
Indira Thillainadarajah
Christina Baghdikian

Agnes Forgacs
Jill Franzosa
Chantel Nicolas
Bhavesh Ahir
Nisha Sipes
Lisa Truong
Max Leung
Kamel Mansouri
Eric Watt
Corey Strope

EPA Collaborators
Kathie Dionisio
Kristin Isaacs
Peter Egeghy
David Dix
Alan Dixon
Scott Lynn
Patience Brown
Don Bergfelt
Les Touart
Rocky Goldsmith

NIH/NCATS
Menghang Xia
Ruili Huang
Anton Simeonov

Others
Alicia Frame (Dow)
Alan Liddell (NCSU)

NTP
Warren Casey
Nicole Kleinstreuer
Mike Devito
Dan Zang
Ray Tice

CERAPP

DTU/food: Technical University of Denmark/ National Food Institute
EPA/NCCT: U.S. Environmental Protection Agency / National Center for Computational Toxicology
FDA/NCTR/DBB: U.S. FDA / National Center for Toxicological Research/Division of Bioinformatics and Biostatistics
FDA/NCTR/DSB: U.S. FDA / National Center for Toxicological Research/Division of Systems Biology
ILS&EPA/NCCT: ILS Inc & EPA/NCCT
IRCSS: Istituto di Ricerche Farmacologiche “Mario Negri”
JRC_Ispra : Joint Research Centre of the European Commission, Ispra.
LockheedMartin&EPA : Lockheed Martin IS&GS/ High Performance Computing
NIH/NCATS : National Institutes of Health/ National Center for Advancing Translational Sciences
NIH/NCI : National Institutes of Health/ National Cancer Institute
RIFM : Research Institute for Fragrance Materials, Inc
UMEA/Chemistry: University of UMEA/ Chemistry department
UNC/MML: University of North Carolina/ Laboratory for Molecular Modeling
UniBA/Pharma: University of Bari/ Department of Pharmacy
UNIMIB/Michem: University of Milano-Bicocca/ Milano Chemometrics and QSAR Research Group
UNISTRA/Infochim: University of Strasbourg/ ChemoInformatique
Helmholtz/ISB: Helmholtz Zentrum Muenchen/Institute of Structural Biology