

In Vitro to In Vivo Extrapolation for High Throughput Prioritization and Decision Making

Workshop Background and Summary of Webinars

John Wambaugh

National Center for Computational Toxicology

Office of Research and Development

U.S. Environmental Protection Agency

wambaugh.john@epa.gov



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

Introduction

- Toxicokinetics (TK) provides a bridge between toxicity and exposure assessment by predicting tissue concentrations due to exposure
 - Traditional TK methods are resource intensive

- Relatively high throughput TK (HTTK) methods have been used by the pharmaceutical industry to determine range of efficacious doses and to prospectively evaluate success of planned clinical trials (Jamei, *et al.*, 2009; Wang, 2010)
 - A key application of HTTK has been “reverse dosimetry” (also called Reverse TK or RTK)

 - RTK can approximately convert *in vitro* HTS results to daily doses needed to produce similar levels in a human for comparison to exposure data

Workshop Webinars

<http://ntp.niehs.nih.gov/pubhealth/evalatm/3rs-meetings/ivive-2016/ivive-2016.html>

- Setting the Stage: Purpose, Definitions, Scope, and Assumptions
Barbara Wetmore
- Building Fit-for-purpose Pharmacokinetic Models
John Wambaugh
- The Role of Pharmacokinetic Model Evaluation
Lisa Sweeney
- Framework for Establishing an Internal Threshold of Toxicological Concern
Corie Ellison

Setting the Stage: Purpose, Definitions, Scope, and Assumptions

***In Vitro - In Vivo* Extrapolation (IVIVE)**

Definition: Utilization of *in vitro* experimental data to predict phenomena *in vivo*

- IVIVE-PK/TK (Pharmacokinetics/Toxicokinetics):
 - Fate of molecules/chemicals in body
 - Considers absorption, distribution, metabolism, excretion (ADME)
 - Uses empirical PK and physiologically-based (PBPK) modeling
- IVIVE-PD/TD (Pharmacodynamics/Toxicodynamics):
 - Effect of molecules/chemicals at biological target *in vivo*
 - Assay design/selection important
 - Perturbation as adverse/therapeutic effect, reversible/ irreversible
- Both contribute to predict *in vivo* effects

Setting the Stage: Purpose, Definitions, Scope, and Assumptions

First Webinar:

- Use of IVIVE tools to incorporate dosimetry has enhanced risk-based interpretation of HTS data
 - As we incorporate more information about chemical kinetics we decrease the uncertainties in high-throughput risk assessment
- Current *in vitro* – *in vivo* assessments for environmental chemicals point to need for tools trained against relevant space for prediction refinement
- IVIVE effort to evaluate PK variability in a manner that could
 1. identify sensitive populations
 2. replace use of default safety factors in risk assessment
- Using IVIVE in PD/TD will require additional considerations to understand chemical concentration at target.

Building Fit-for-purpose Pharmacokinetic Models

Second Webinar:

- We must keep in mind the purpose – simple models appear to allow meaningful prioritization of further research
- A primary application of HTKK is “Reverse Dosimetry” or RTK
 - Can infer daily doses that produce plasma concentrations equivalent to the bioactive concentrations,
- We can also use QSAR to build provisional PBTK models

But we must consider parsimony and domain of applicability:

- Do not build beyond the ability to evaluate predictions
- Carefully determine whether, when, and why model errors are conservative
- Collect PK data from *in vivo* studies to allow larger, systematic studies
- R package “httk” freely available on CRAN allows statistical analyses

The Role of Pharmacokinetic Model Evaluation

Third webinar:

- Model evaluation principles are applicable to models of varying complexity
- Model evaluation is dependent on having a context for model use/application
- Formal sensitivity analysis can focus model evaluation on key parameters
- Even “simple” models can be challenging to evaluate
- In general, there are good reasons to believe the human H₁TK models being generated for IVIVE are sufficiently accurate for the intended application
 - The tendency for these models to err in a conservative direction may not be a significant drawback in that context

Framework for Establishing an Internal Threshold of Toxicological Concern

Fourth webinar:

- Registrants attempted to use metabolism based read-across to support their chemical
 - Parent half life in blood ~ 15 minutes
 - PBPK modeling demonstrated that parent AUC was <1% of metabolite AUC following exposure to parent chemical (i.e. predominant systemic exposure is to metabolite) threshold.
- Registrants were unable to adequately justify why the low level, short term systemic exposure to the parent would not represent human safety concern. As such, they had to perform a developmental toxicity study in rodents.
- Availability of an internal TTC may have allowed for comparison of the systemic exposure to an internal exposure threshold.

Published Case Study:



The challenge of using read-across within the EU REACH regulatory framework; how much uncertainty is too much? Dipropylene glycol methyl ether acetate, an exemplary case study



Nicholas Ball^{a,*}, Michael Bartels^b, Robert Budinsky^b, Joanna Klapacz^b, Sean Hays^c, Christopher Kirman^d, Grace Patlewicz^e

^a Toxicology & Environmental Research and Consulting (TERC), Dow Europe GmbH, Bachtobelstrasse 3, CH-8810 Horgen, Switzerland

^b Toxicology & Environmental Research and Consulting (TERC), The Dow Chemical Company, 1803 Bldg, Washington Street, Midland, MI 48674, USA

^c Summit Toxicology, 165 Valley Road, Lyons, CO 80540, USA

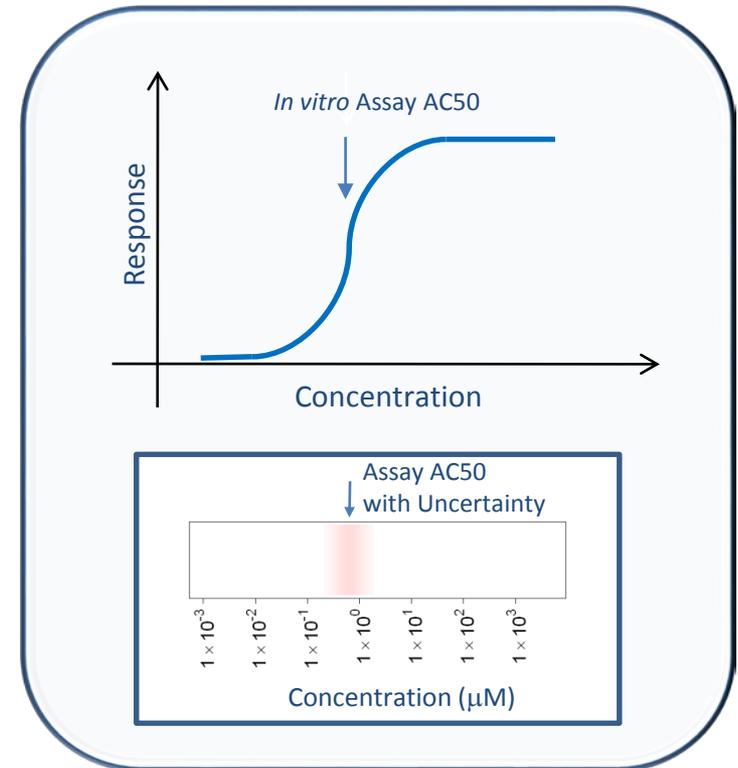
^d Summit Toxicology, 29449 Pike Drive Orange Village, OH 44022, USA

^e DuPont Haskell Global Centers for Health and Environmental Sciences, 1090 Elkton Road, Newark, DE 19711, USA

Building Fit-for-purpose Pharmacokinetic Models

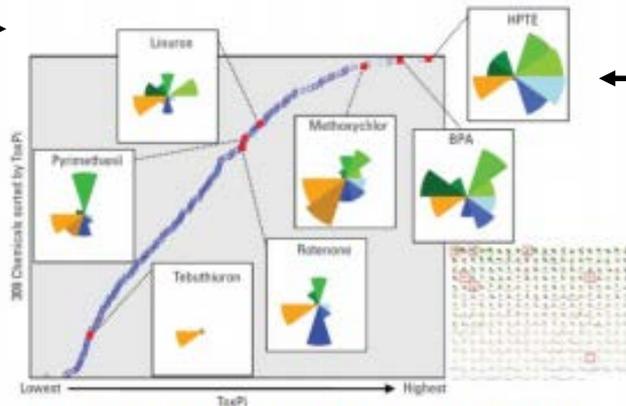
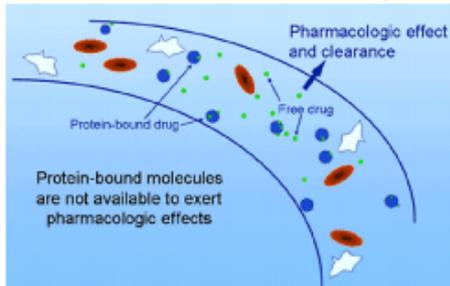
High Throughput Bioactivity

- **Tox21:** Examining >10,000 chemicals using ~50 assays intended to identify interactions with biological pathways (Schmidt, 2009)
- **ToxCast:** For a subset (>1000) of Tox21 chemicals ran >800 additional assay endpoints (Judson et al., 2010)
- Most assays conducted in dose-response format (identify 50% activity concentration – AC50 – and efficacy if data described by a Hill function)
- All data are public: <http://actor.epa.gov/>



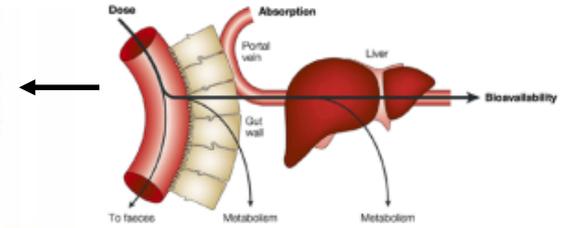
Setting the Stage: Purpose, Definitions, Scope, and Assumptions

Protein Binding



Reif et al. Environ Hlth Perspect 118:1714, 2010

Bioavailability



van de Waterbeemd and Gifford, *Nat Rev Drug Disc* 2:192, 2003

Prioritization and hazard prediction based on nominal (*in vitro*) concentrations can misrepresent potential health risks

Metabolic Clearance



Use *in vitro-in vivo* extrapolation (IVIVE)

The Role of Pharmacokinetic Model Evaluation

- Dose-response relationships can be divided into pharmacokinetic (PK) and pharmacodynamic (PD) aspects
 - PK: “what the body does to the chemical”
 - PD: “what the chemical does to the body”
- Traditional PK/TK studies are resource intensive
- PK and PD data and models are important in risk assessment they connect exposure and toxicity



Framework for Establishing an Internal Threshold of Toxicological Concern

PK Modeling Approaches:

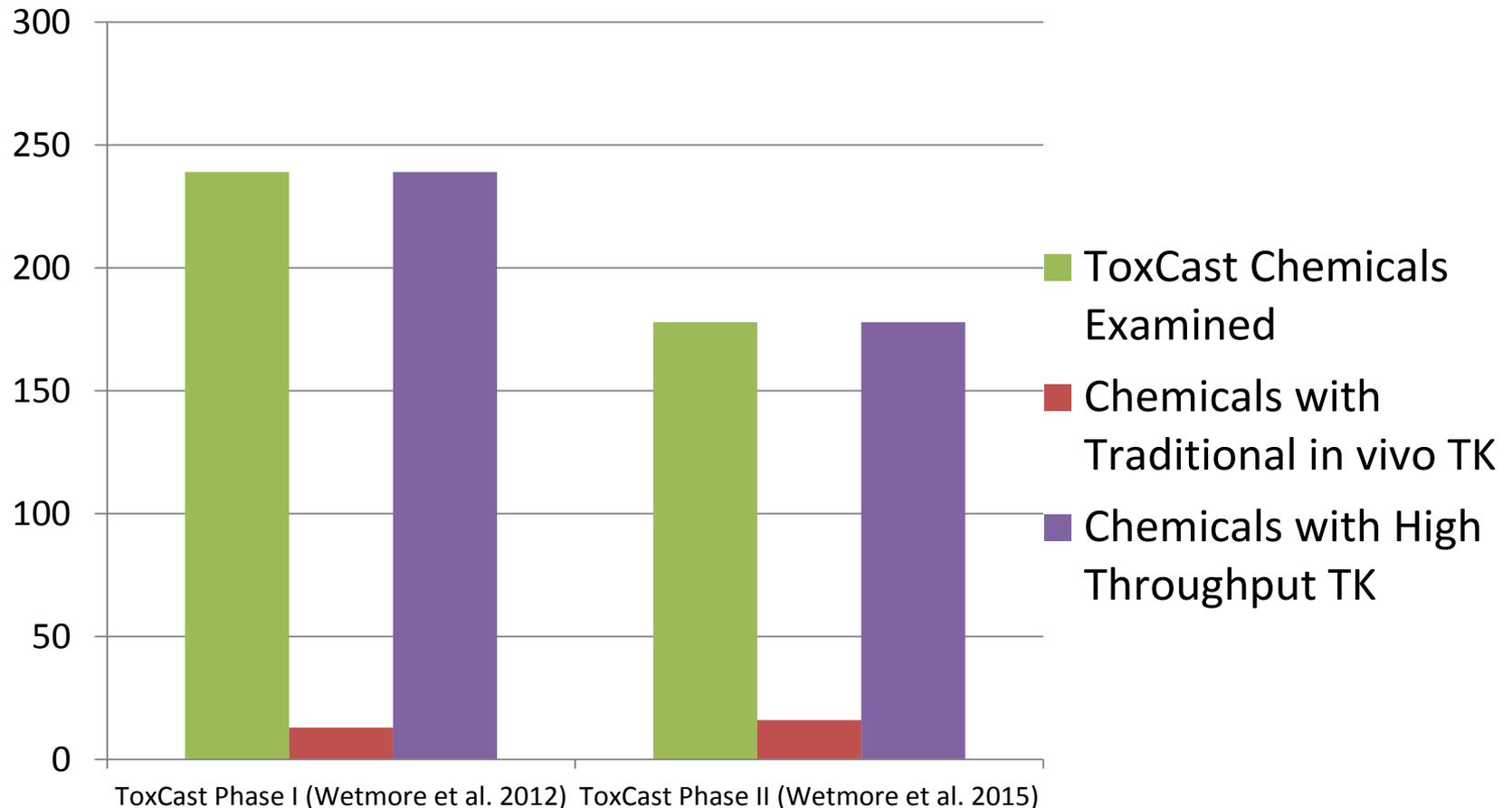
- Multiple pharmacokinetic approaches available as options to use in framework:

- C_{ss} equation
$$C_{ss} = \frac{k_0 \times F}{(GFR \times F_{ub}) + \left[\frac{(Q_l \times F_{ub} \times Cl_{int})}{(Q_l + F_{ub} \times Cl_{int})} \right]}$$
 Wilkinson and Shand (1975)

- Commercially available generic PBPK models
 - GastroPlus (Simulations plus)
 - ADME WorkBench (Aegis Technologies)
 - SimCyp
- Freely available generic PBPK models

Building Fit-for-purpose Pharmacokinetic Models

The need for higher throughput *in vitro* toxicokinetics

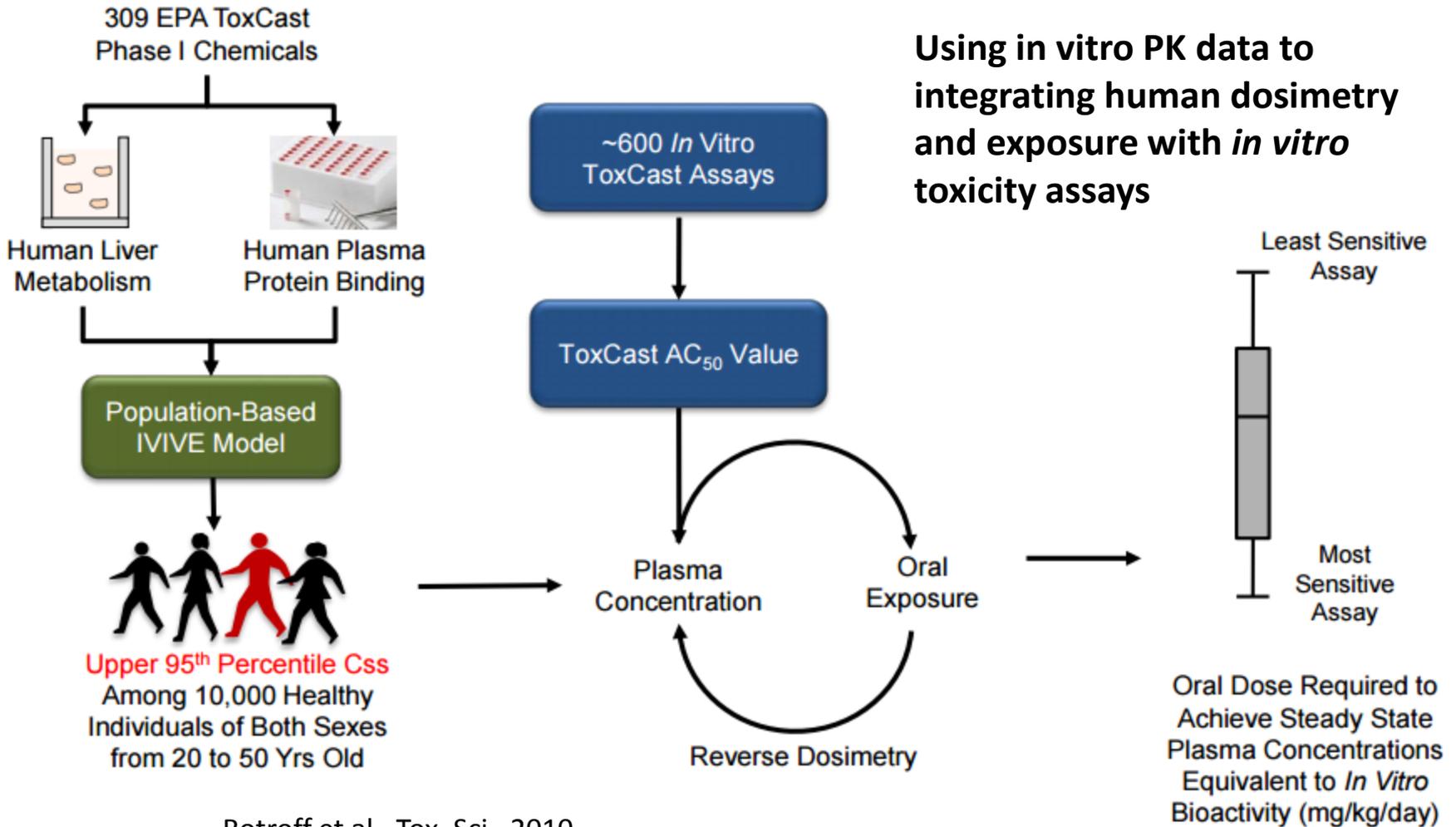


Framework for Establishing an Internal Threshold of Toxicological Concern

In Silico Prediction of Parameters:

- Various options for predicting ADME parameters
 - Swiss Institute of Bioinformatics provides summary of software, web services & databases
 - <http://www.click2drug.org/index.html>
 - Multiple published algorithms for different ADME input parameters
- Robust in silico approaches for predicting metabolism are not currently available
 - Quantitative Structure-Activity Relationships (QSARs) developed to date have limited applicability domain

Setting the Stage: Purpose, Definitions, Scope, and Assumptions



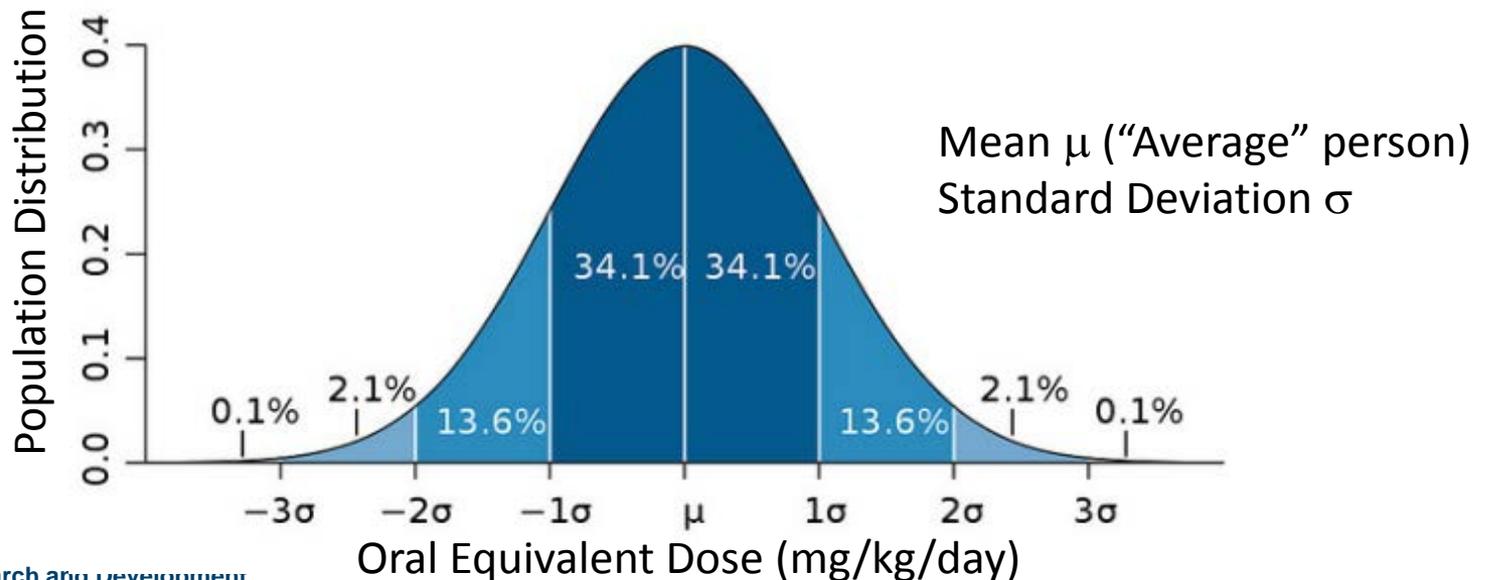
Using *in vitro* PK data to integrating human dosimetry and exposure with *in vitro* toxicity assays

Rotroff et al., Tox. Sci., 2010
Wetmore et al., Tox Sci., 2012
Wetmore et al., Tox Sci., 2015

Slide from Barbara Wetmore's webinar

The Role of Pharmacokinetic Model Evaluation

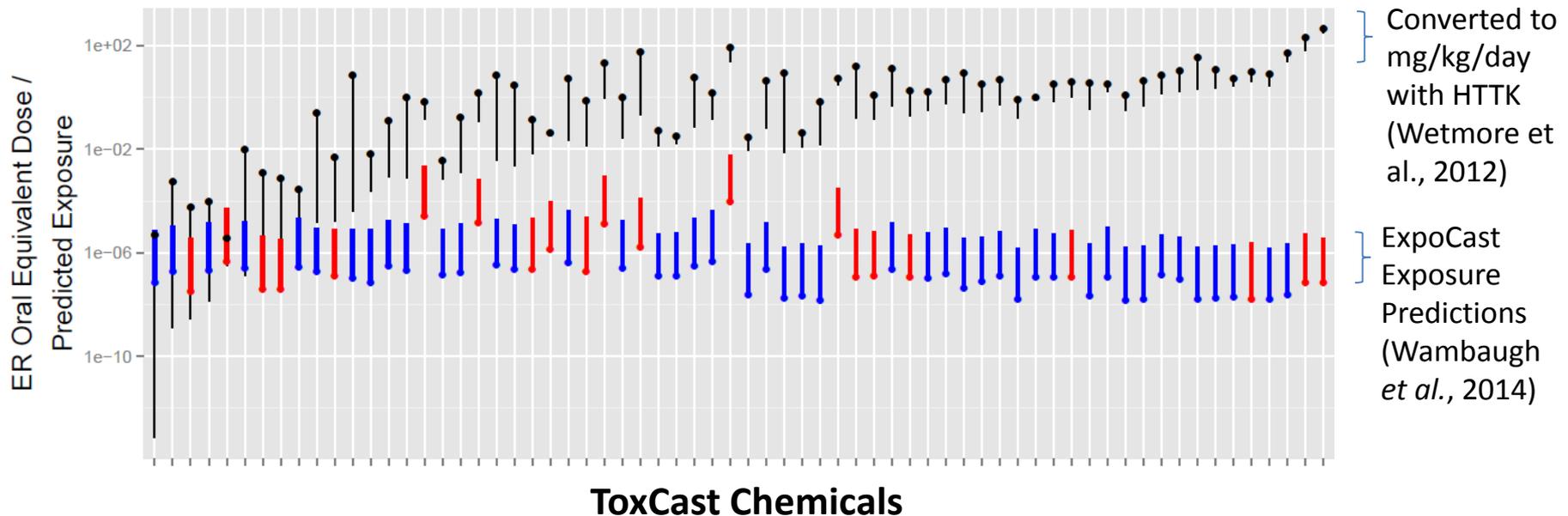
- Simplistic models are used to estimate oral equivalent dose (OED) for an effective in vitro concentration
 - *E.g.*, dose that in 95% of simulated individuals produces steady-state blood concentrations below the lowest effective in vitro concentration
- OEDs are compared to exposure estimates to prioritize chemicals for research/testing



Building Fit-for-purpose Pharmacokinetic Models

Pharmacokinetics allows context for high throughput screening data

Endocrine disruption AOP (Judson et al., in prep.)



December, 2014 Panel:

“Scientific Issues Associated with Integrated Endocrine Bioactivity and Exposure-Based Prioritization and Screening”

The Role of Pharmacokinetic Model Evaluation

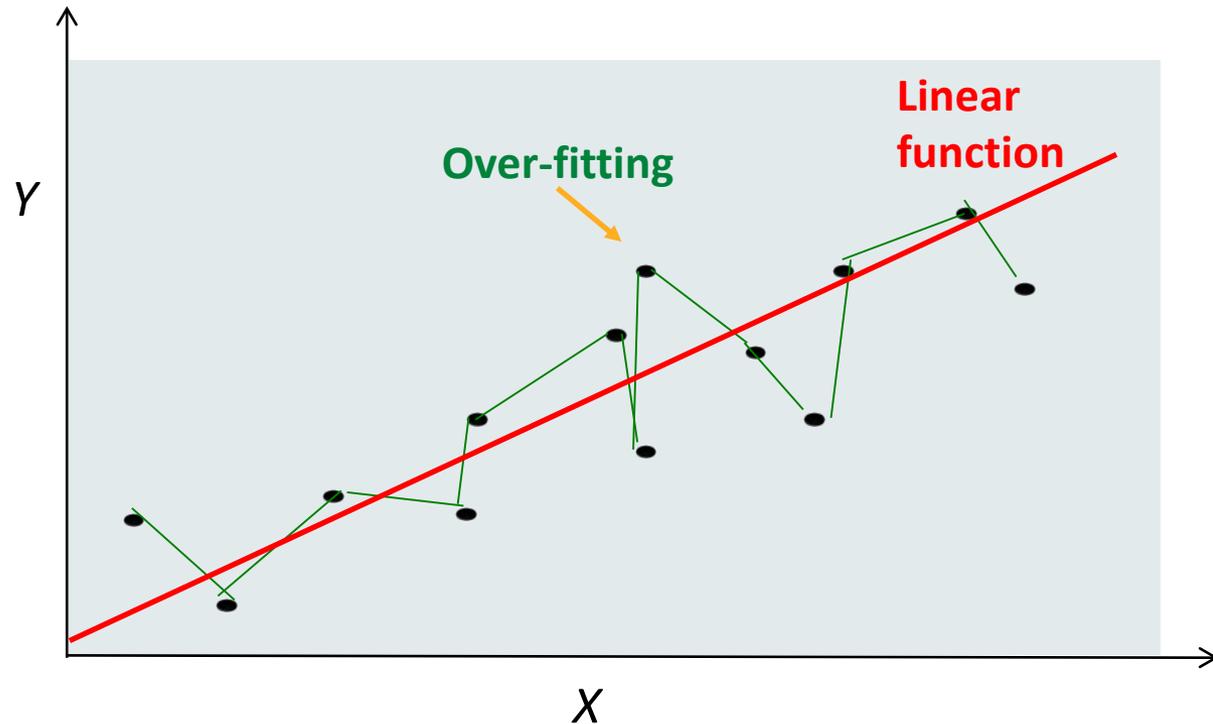
- **Goal: To assess model confidence for either a specific application or a spectrum of (tiered) applications**
 - Prioritization vs. IRIS RfD or slope factor
 - Level of model confidence vs. acceptable margin of exposure
- We will assume a model has already been built
 - Model building is frequently iterative
 - Initial model evaluation may identify modifications required/desired for a particular purpose
- Key questions adapted from McLanahan et al. (2012)

Building Fit-for-purpose Pharmacokinetic Models

Lex Parsimoniae “Law of Parsimony”

“Among competing hypotheses, the one with the fewest assumptions should be selected.” William of Ockham

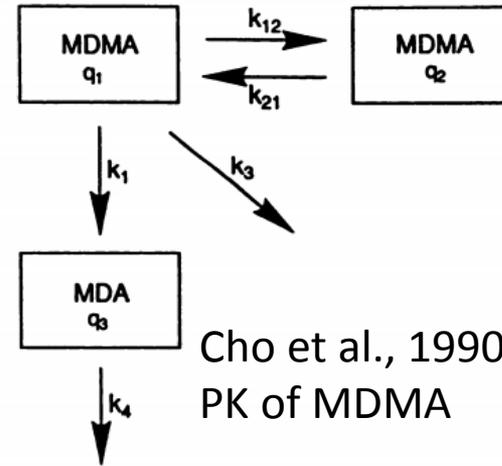
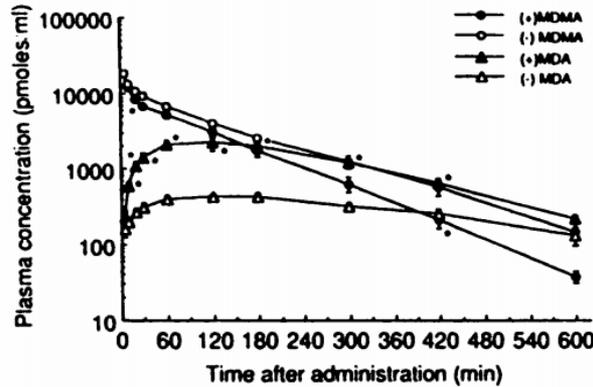
“As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality.” Albert Einstein



Building Fit-for-purpose Pharmacokinetic Models

Complexity should fit the data...

“Since all models are wrong the scientist cannot obtain a ‘correct’ one by excessive elaboration. On the contrary[,] following William of Occam[,] they] should seek an economical description of natural phenomena.” George Box, University of Wisconsin



Jones et al., 2012
PK of Statins

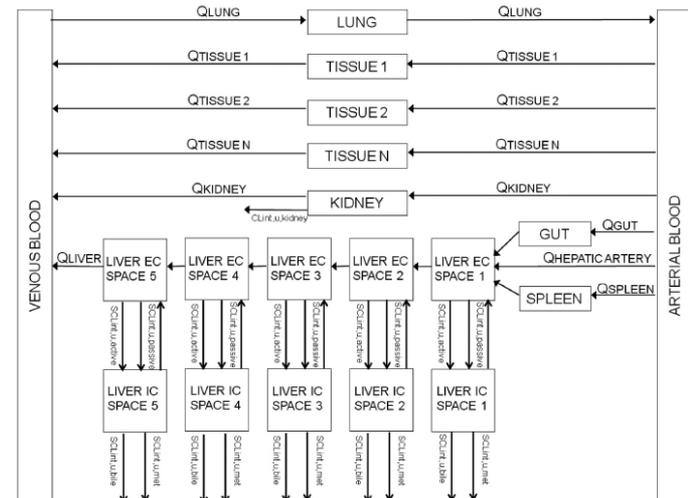


FIG. 2. Schematic diagram of the in vivo PBPK model. EC, extracellular, IC, intracellular.

The Role of Pharmacokinetic Model Evaluation



Key Questions:

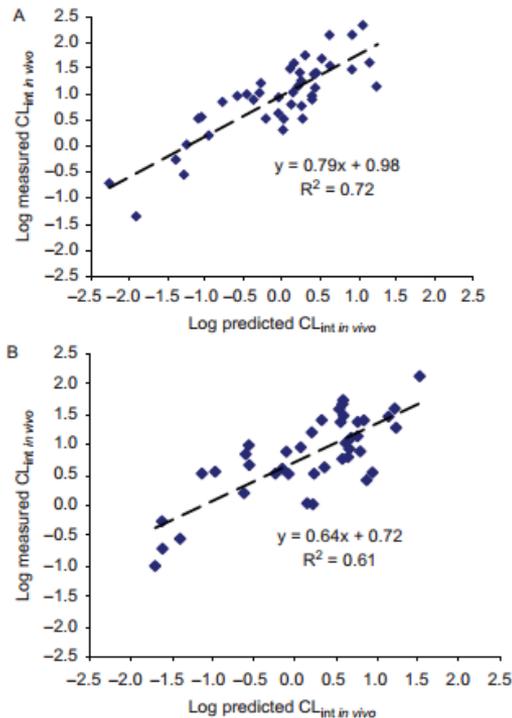
- Is the model verifiable?
 - Can previous simulations be reproduced?
- Evaluate model performance
 - Has model been tested against all (or most) of the appropriate literature data?
 - Not all published models have been comprehensively evaluated
 - How well did the model perform?
 - How good is “good enough”?
 - One recommendation is, on average, within a factor of 2 (IPCS, 2010)
 - How well is the model known/expected to perform in the scenario of interest (e.g., low vs. high concentrations)

Building Fit-for-purpose Pharmacokinetic Models

Characterizing Accuracy of HTTK – Wang (2010): *In vitro* predictions typically within a factor of three for pharmaceuticals

Pharmaceuticals:

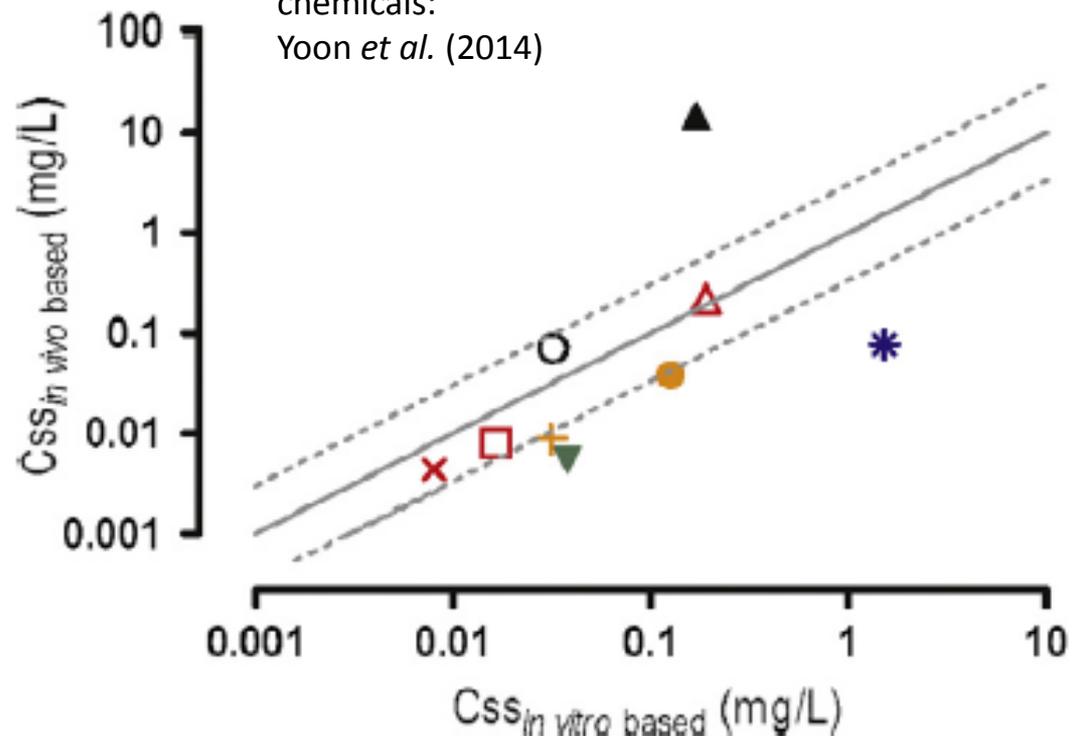
Sohlenius-Sternbeck *et al.* (2010)



Environmental

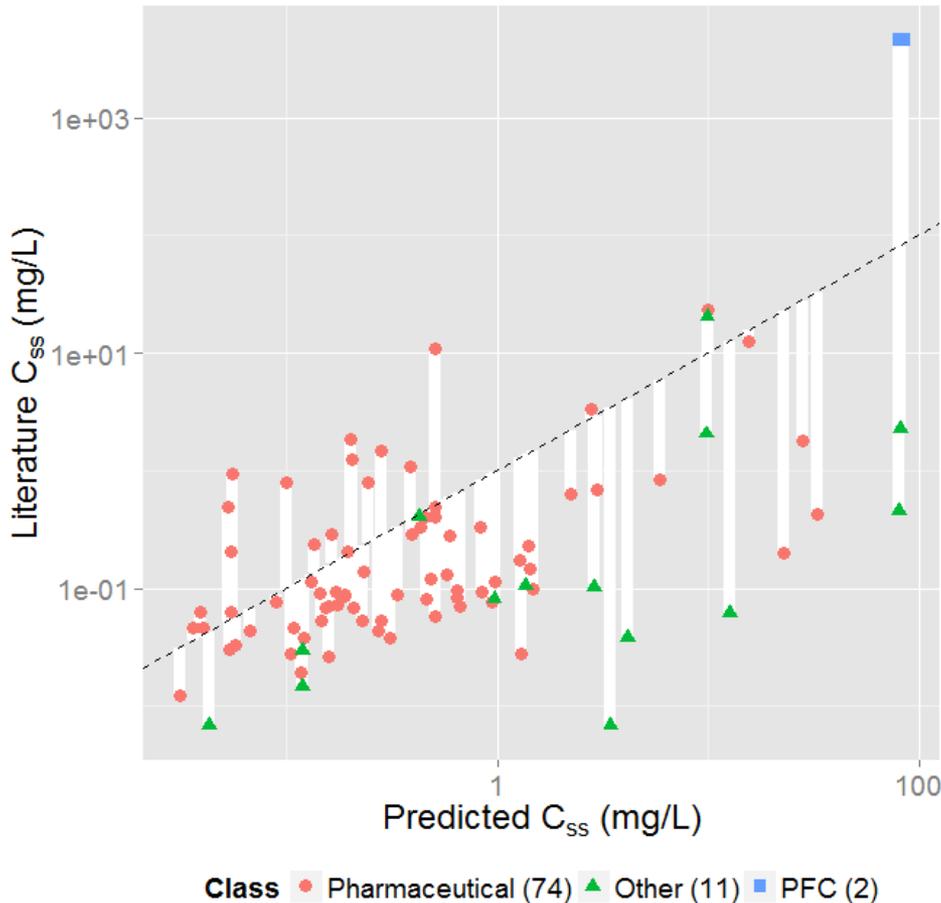
chemicals:

Yoon *et al.* (2014)



Building Fit-for-purpose Pharmacokinetic Models

Using *in vivo* data to evaluate HTKK

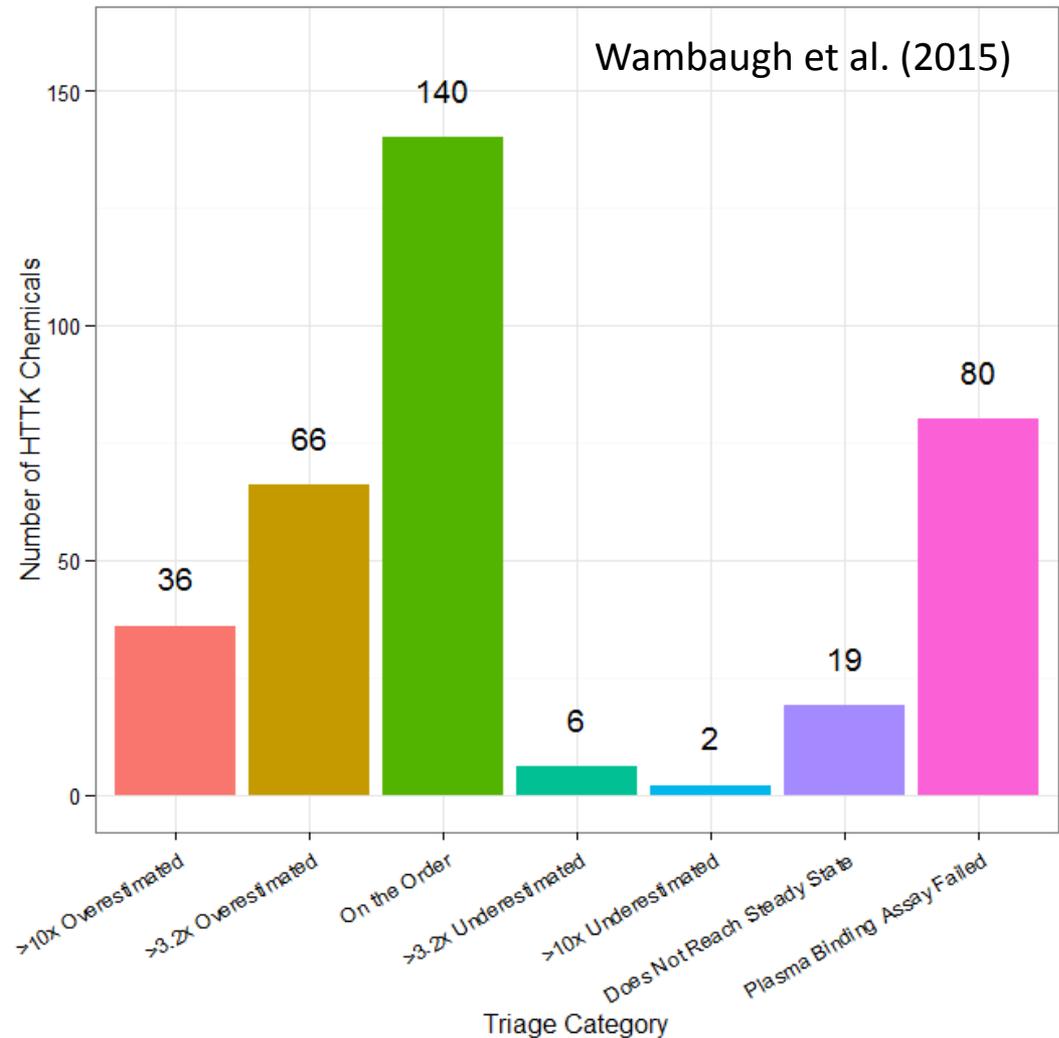


- When we compare the C_{ss} predicted from *in vitro* HTKK with *in vivo* C_{ss} values determined from the literature we find limited correlation ($R^2 \sim 0.34$)
- The dashed line indicates the identity (perfect predictor) line:
 - Over-predict for 65
 - Under-predict for 22
- The white lines indicate the discrepancy between measured and predicted values (the residual)

Building Fit-for-purpose Pharmacokinetic Models

Toxicokinetic Triage

- Through comparison to *in vivo* data, a cross-validated (random forest) predictor of success or failure of HTTK has been constructed
- Add categories for chemicals that do not reach steady-state or for which plasma binding assay fails
- All chemicals can be placed into one of seven confidence categories



Setting the Stage: Purpose, Definitions, Scope, and Assumptions

Reasons for C_{ss} Over-prediction - Opportunities for Refinement

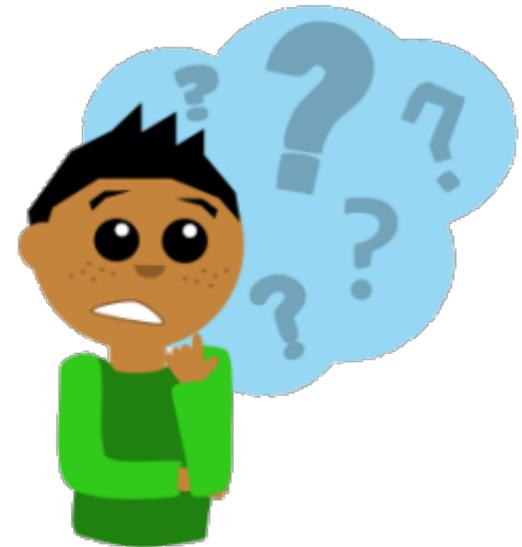
- Not all routes of metabolic clearance are captured
 - Extrahepatic (intestinal, renal, etc.) metabolism
 - Non-hepatocyte-mediated clearance
- Hepatocyte suspensions unable to detect clearance of low turnover compounds
- Absorption / Bioavailability assumed 100%
- Restrictive vs. Nonrestrictive clearance
- Conservative assumptions drive poor predictive ability for chemicals known to be rapidly cleared *in vivo*

The Role of Pharmacokinetic Model Evaluation

Key Questions:

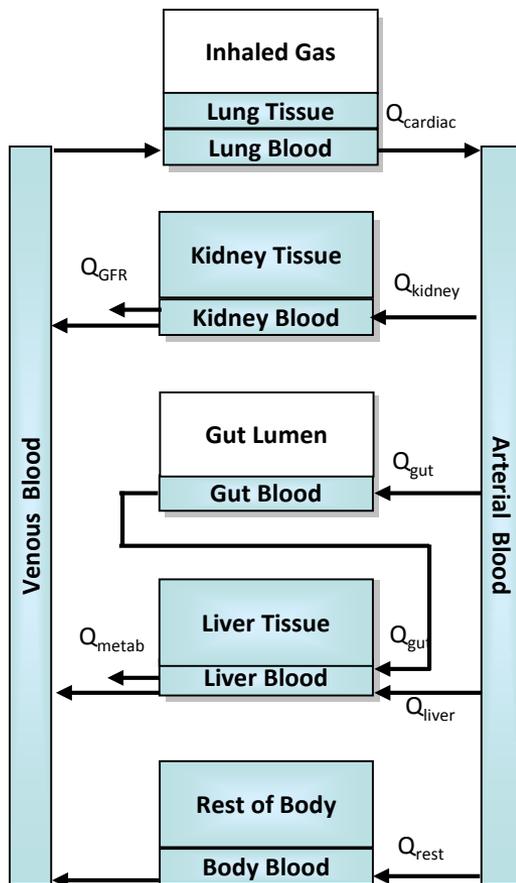
-  • How biologically realistic is the model structure vs. how realistic does it need to be?
 - Lumping vs. splitting

-  • Is the model suitable for intended use? For what uses is the model suitable?
 - Species, exposure route/scenario, suitable metrics
 - Simplified, steady-state models may not be suitable for short, dynamic life stages (e.g. pregnancy)



Building Fit-for-purpose Pharmacokinetic Models

A general physiologically-based pharmacokinetics (PBPK) model



“httk” R Package

<https://cran.r-project.org/web/packages/httk/>

Can access this from the R GUI: “Packages” then “Install Packages”

543 Chemicals to date

443 PBPK models

More data being collected, analyzed, and published on a regular basis

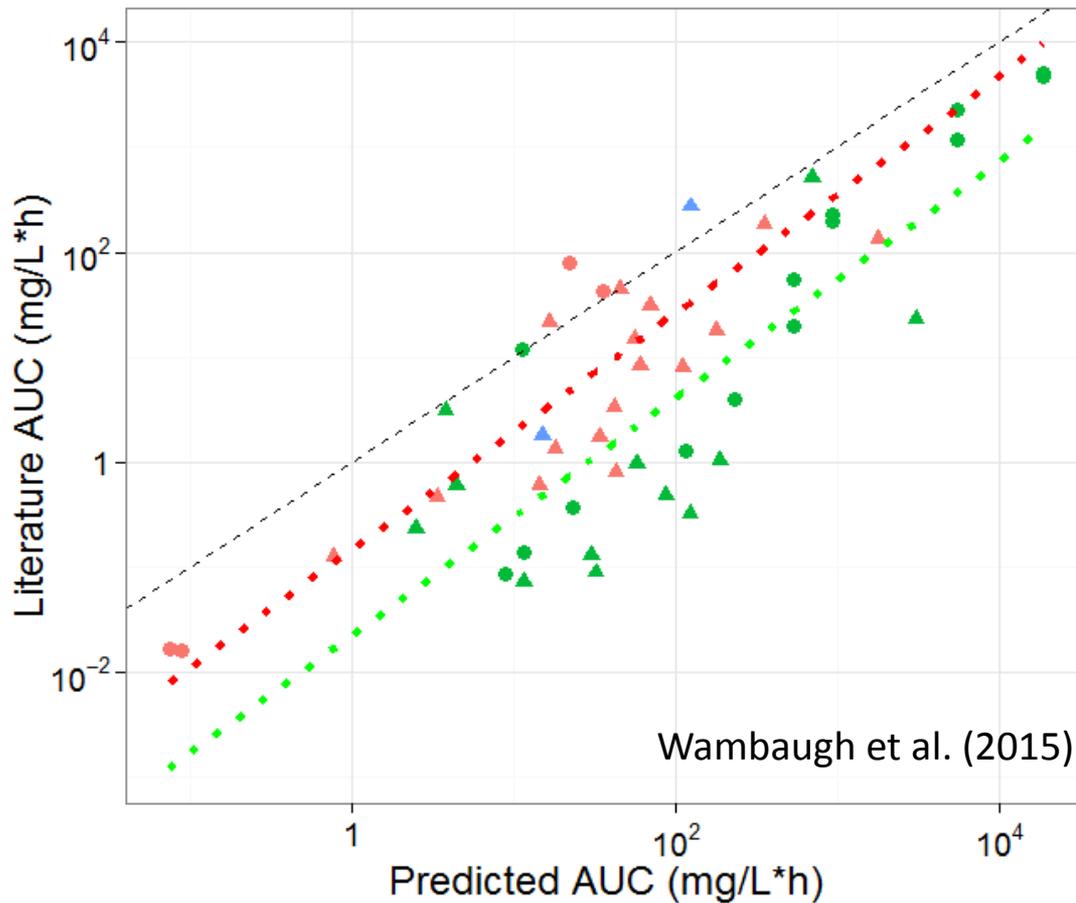
Pearce *et al.* accepted at Journal of Statistical Software

Can test impact of assumptions such as lumping vs. splitting systematically

See poster by Robert Pearce

Building Fit-for-purpose Pharmacokinetic Models

A general physiologically-based pharmacokinetics (PBPK) model



- HTPBPK predictions for the AUC (time integrated plasma concentration or Area Under the Curve)
- *in vivo* measurements from the literature for various treatments (dose and route) of rat.
- Predictions are generally conservative – *i.e.*, predicted AUC higher than measured
- Oral dose AUC ~6.4x higher than intravenous dose AUC

See poster by Nisha Sipes

Slide from John Wambaugh's webinar

Building Fit-for-purpose Pharmacokinetic Models

New In Vivo PK Data Set

- Could the difference be related to inhomogeneous C_{ss} data?
 - Initially relying on Obach (2008) data plus data curated by TNO (Sieto Bosgra lead) from literature
- Only 13 non-pharmaceuticals examined so far
- Cross lab study:
 - 20 chemicals examined by NHEERL (Mike Hughes lead)
 - 8 chemicals examined by RTI (Tim Fennell lead)
 - 2 overlap chemicals (Bensulide and Propyzamide)
 - See poster by Mike Hughes

See poster by Michael Hughes

Slide from John Wambaugh's webinar

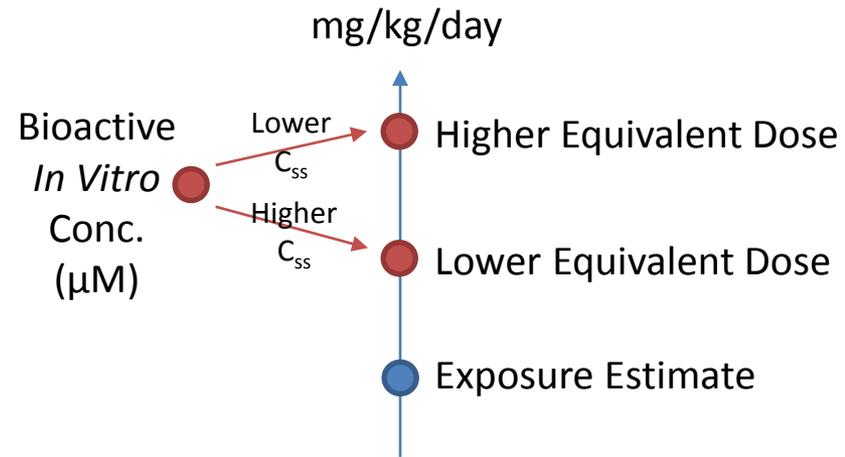
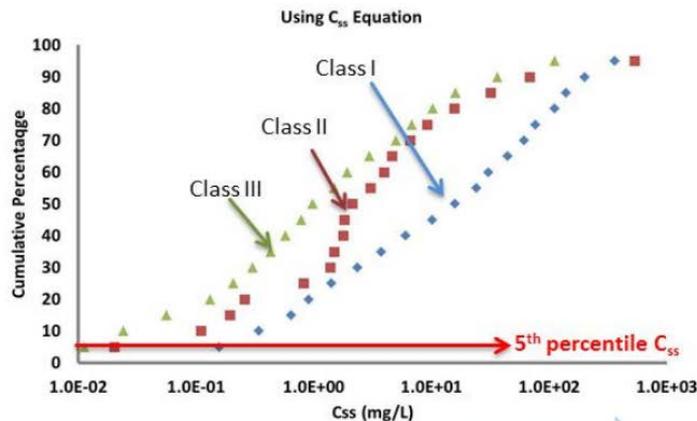
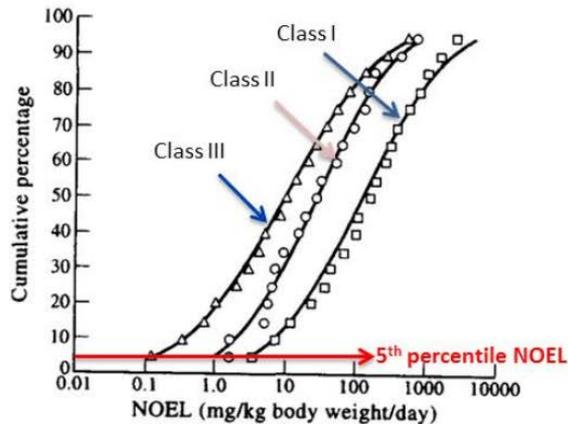
Framework for Establishing an Internal Threshold of Toxicological Concern

Assessing the Impact of IVIVE Bias

Reverse dosimetry:

Overestimations by IVIVE leads to lower predicted ratios between bioactive doses and expected exposures, thus leading to a higher predicted risk

Forward dosimetry:



Overestimations by IVIVE leads to higher estimated tissue concentrations linked to toxic effect – **underestimating potency**

- Toxicokinetics (TK) provides a bridge between hazard and exposure by predicting tissue concentrations due to exposure
 - Higher throughput toxicokinetics (HTTK) appears to provide essential data
- We must keep in mind the purpose – simple models appear to allow meaningful prioritization of further research
- A primary application of HTTK is “Reverse Dosimetry” or RTK
 - We can infer daily doses that produce plasma concentrations equivalent to the bioactive concentrations identified by HTS

But we must consider parsimony and domain of applicability



The horse is out of the barn, these data and models are being used – what are the most necessary refinements and caveats?