United States Environmental Protection Agency

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.

A multi-tiered hierarchical Bayesian approach to derive toxic equivalency factors (TEFs)

Caroline L. Ring, Ph.D

Center for Computational Toxicology and Exposure
United States Environmental Protection Agency

Presented to EPA/NICEATM Workshop: Advancing Quantitative Analysis in Human Health Assessments through Probabilistic Methods

Background: Dioxin-like compounds

(DLCs) are "a group of chemical compounds that share certain chemical structures and biological characteristics" [1]

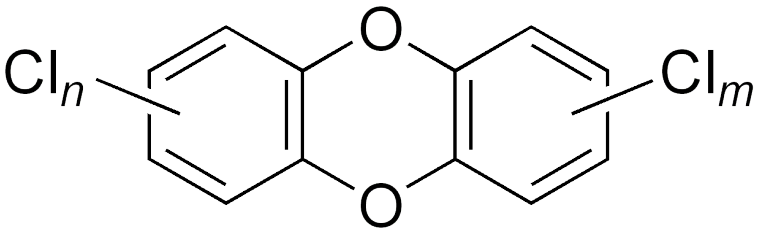
- Polychlorinated dibenzo-*p*-dioxins (PCDDs)
- Polychlorinated dibenzofurans (PCDFs)
- Polychlorinated biphenyls (PCBs)

Usually occur as **mixtures**

29 DLCs exhibit toxicity, via the same mechanism: binding to aryl hydrocarbon (AhR) receptor

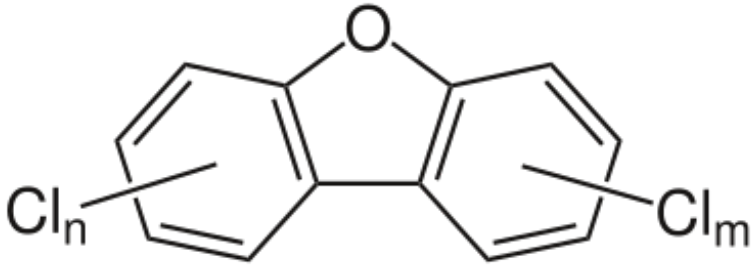
- transcription factor affecting expression of many genes
- Many different adverse biological effects

[1] https://www.epa.gov/dioxin/learn-about-dioxin



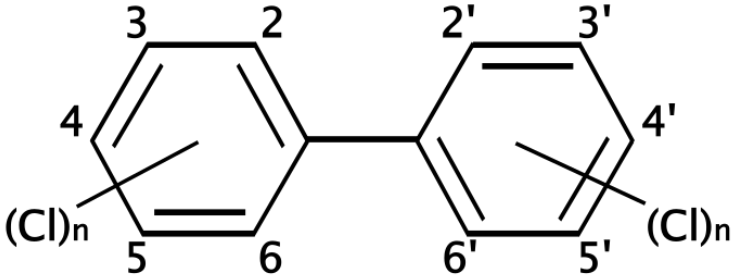
General chemical structure of PCDDs

By Edgar181 - Own work, Public Domain, https://commons.wikimedia.org/w/index.php?curid=5428581



General chemical structure of PCDFs

By Leyo - Own work, Public Domain, https://commons.wikimedia.org/w/index.php?curid=7106630



General chemical structure of PCBs

By D.328 - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=1048994

Toxicity equivalence framework for dioxin-like compounds [EPA, 2010]

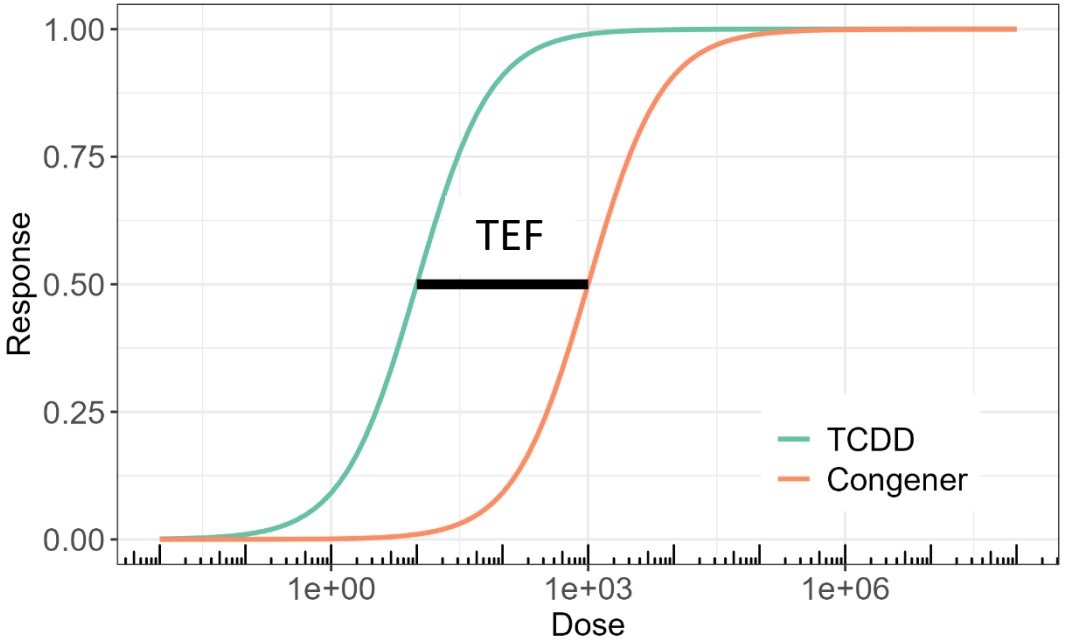
- DLCs exhibit **additive toxicity**
- Toxicity of each congener expressed relative to **index compound**, 2,3,7,8-TCDD

Toxicity equivalence factor (TEF):
"consensus estimates of compound-specific toxicity/potency of a congener, relative to the toxicity/potency of index chemical" (EPA, 2010)

Toxicity equivalence quotient (TEQ):

$$TEQ = \sum_{i=1}^{N_{congeners}} \text{conc}_i \times TEF_i$$

Parallel or similarly-shaped curves (EPA, 2010)



The toxicity equivalence (TEF/TEQ) framework allows rapid estimation of risk from exposure to mixtures of congeners.

TEFs are estimated from studies of **relative potency (REP)**

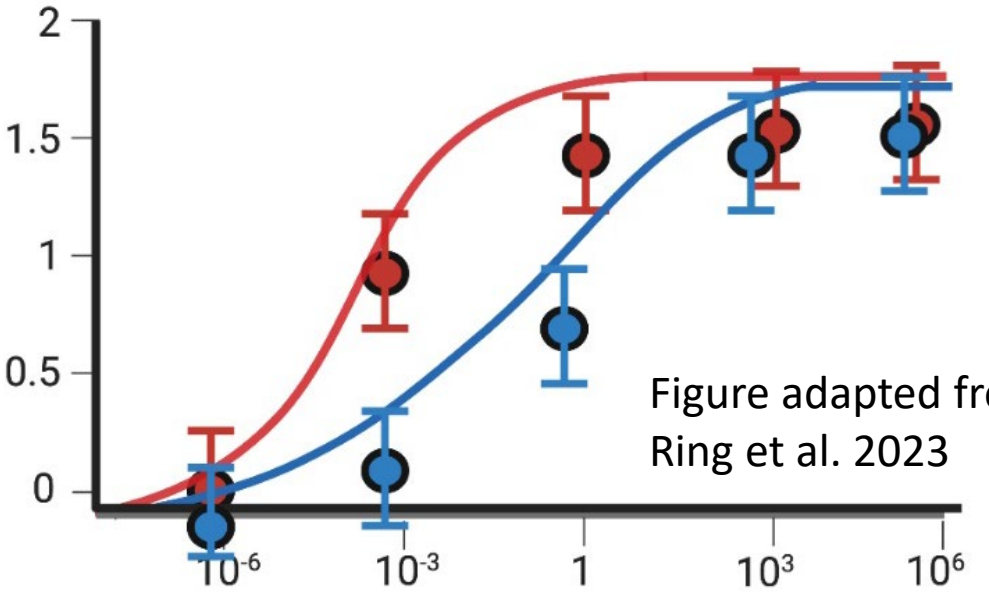


Figure adapted from Ring et al. 2023

$$REP_i = \frac{\text{potency for TCDD}}{\text{potency for congener } i}$$

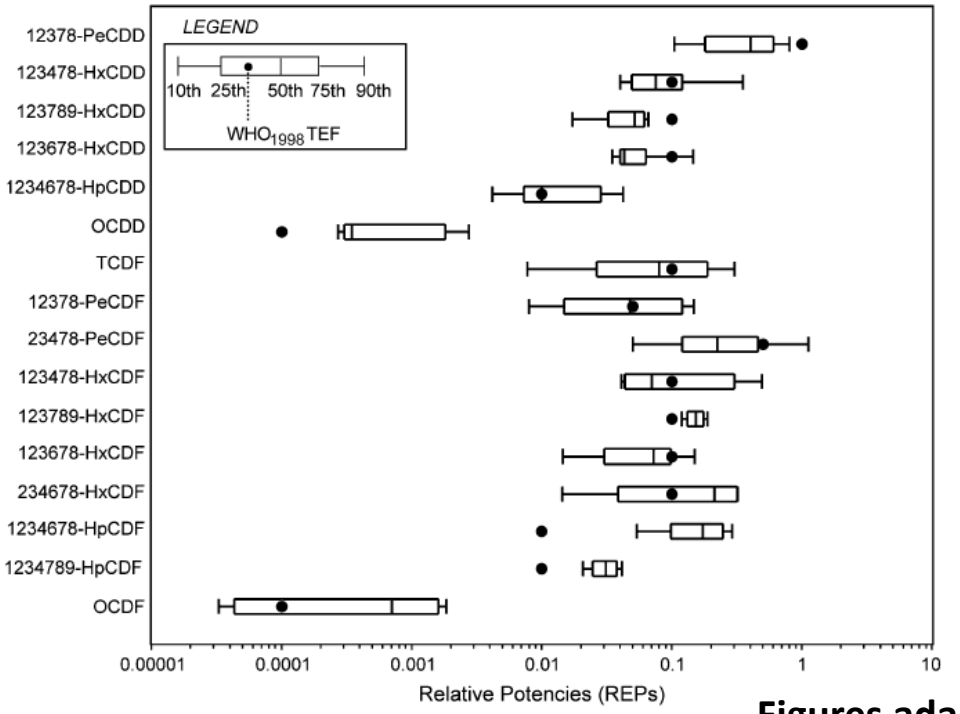
Relative potency can be calculated in different ways (ratio of ED50s, ED20s, BMDs, NOAEL/LOAELs…)

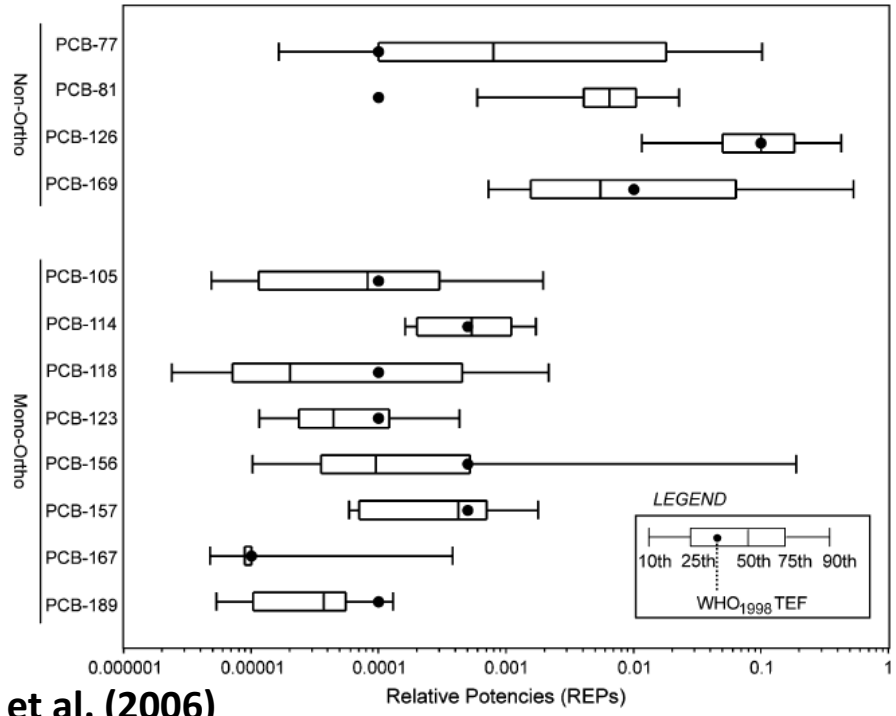
Usually, only point-estimate REP is reported – uncertainty not quantified

In 2005, TEFs were determined by WHO expert panel from qualitative assessment of an evidence base of relative potency studies (REP_{2004})

- 83 publications, 634 REP values
  - Mammals or mammalian cells
  - Both *in vivo* and *in vitro* studies included
  - Wide variety of endpoints (both toxic and biochemical)
- REP distributions were only used as starting points for expert judgment
  - TEFs were *not* chosen as fixed percentiles
- TEFs were assigned in half-log increments (rough uncertainty quantification)

Van den Berg et al (2006)
Haws et al (2006)



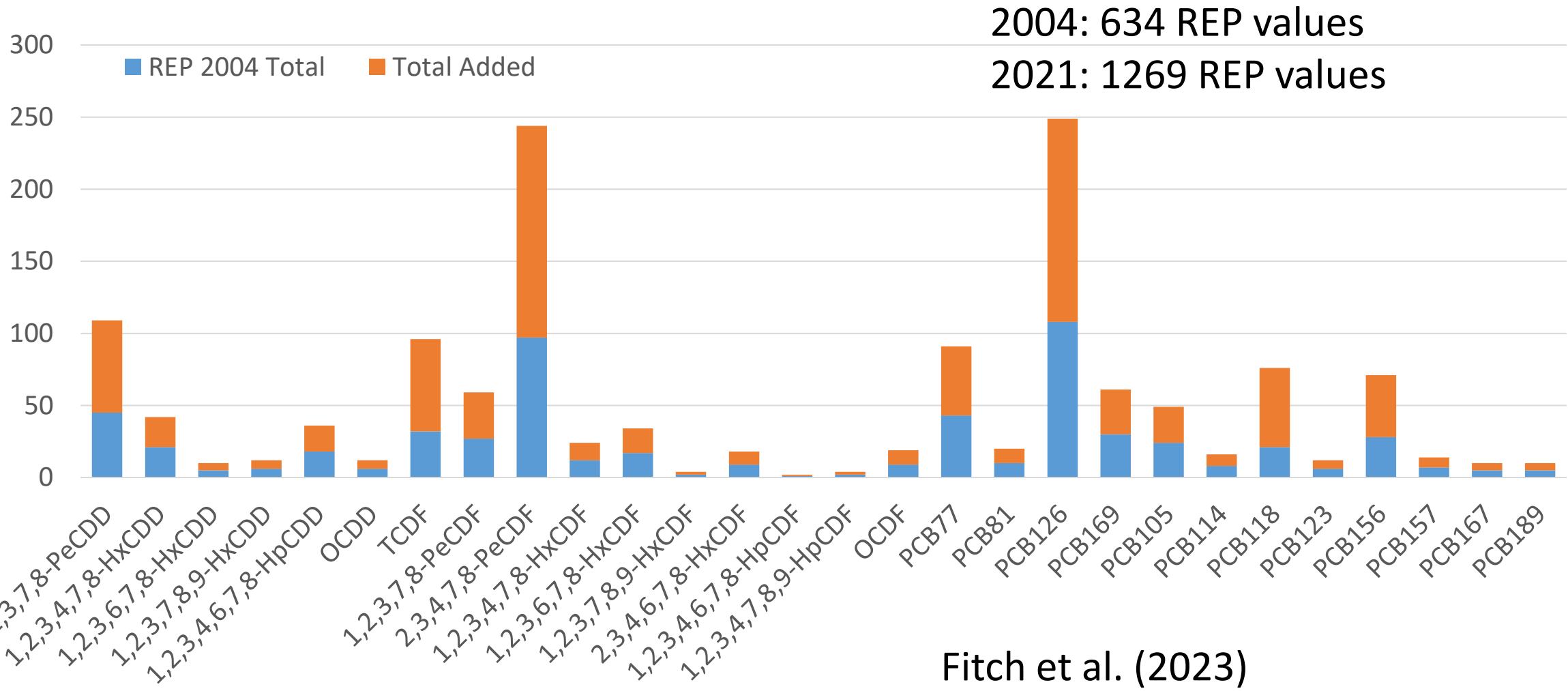


Figures adapted from Haws et al. (2006)

WHO (2005) expert panel noted varying **reliability**, **relevance**, and **amount** of REP data, and the need to weight it accordingly

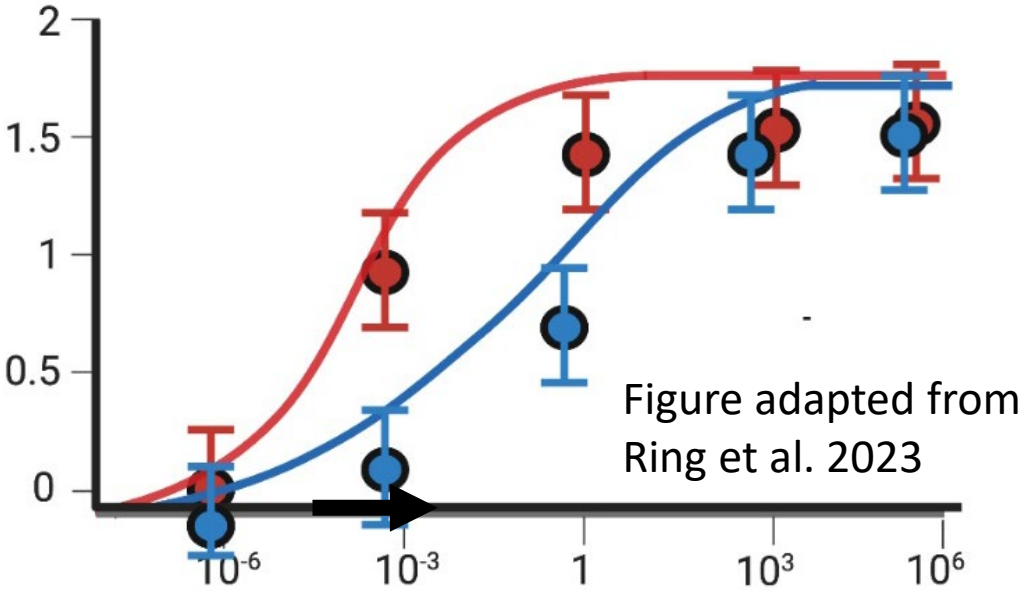
- **REPs measured by higher-quality studies should be more heavily weighted** (van den Berg et al., 2006)
  - E.g., less uncertainty in extrapolating from *in vivo* vs. *in vitro* studies
- **Uncertainty from differing REP calculation methods**
  - What metric of potency was used?
  - Uncertainty in dose-response modeling
  - Were curves parallel?
- **Database uncertainty**: Some congeners have many REP studies; others have few (Haws et al., 2006)
- In 2005, weighting and database uncertainty was handled using **qualitative expert judgment**.
  - Panel recommended developing a **quantitative consensus weighting scheme** in future [Van den Berg et al. (2006), Haws et al. (2006)].

In 2021, database of REP studies was updated



Fitch et al. (2023)

REP_{2021} database also now includes original dose-response data, where available (570 of 1269 REP studies)



Author-derived REP, e.g. 0.001

Dose-response data allows evaluation of the assumption of parallel curves

Dose-response data also allows estimation of uncertainty in each REP

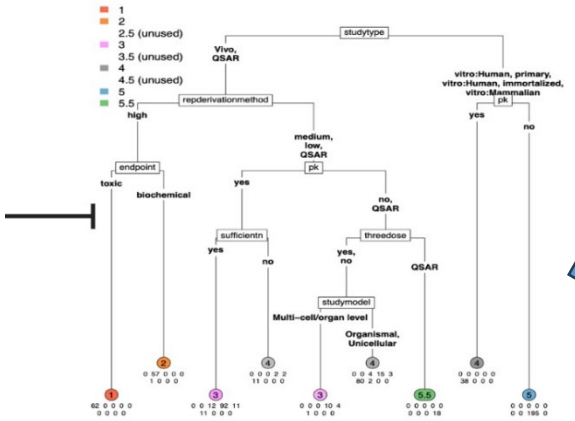
EPA United States Environmental Protection Agency

Updated TEF analysis using REP_{2021} database

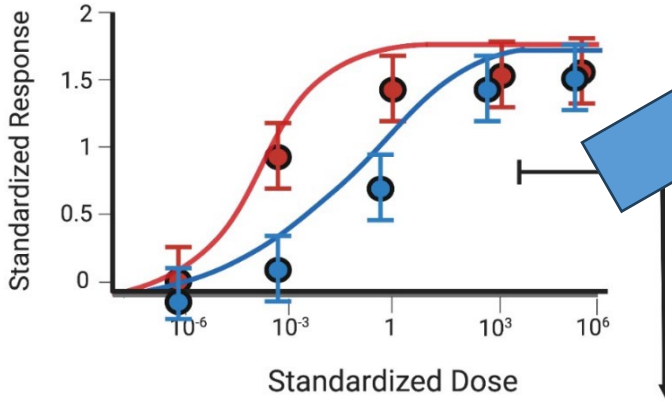
- Transparent & reproducible
- All assumptions made explicit
- Incorporate quantitative weighting based on study quality (reliability & relevance)
- Quantify uncertainty

Best Estimate TEF Workflow

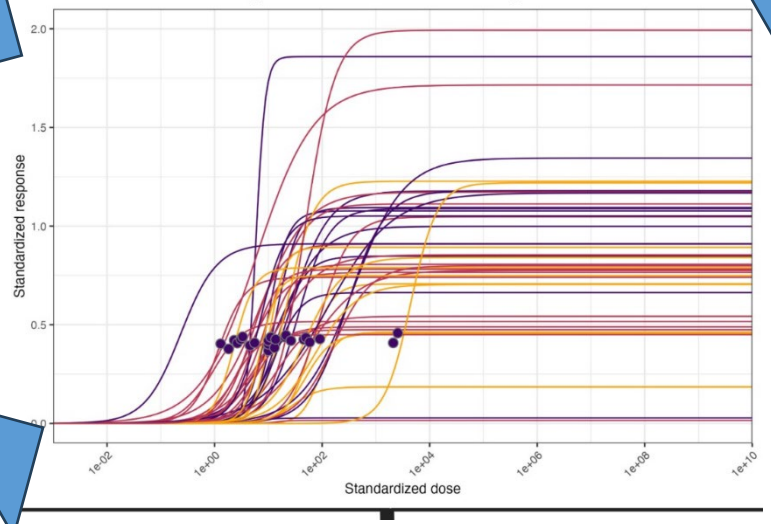
Machine Learning-Based REP Dataset Quality Predictions



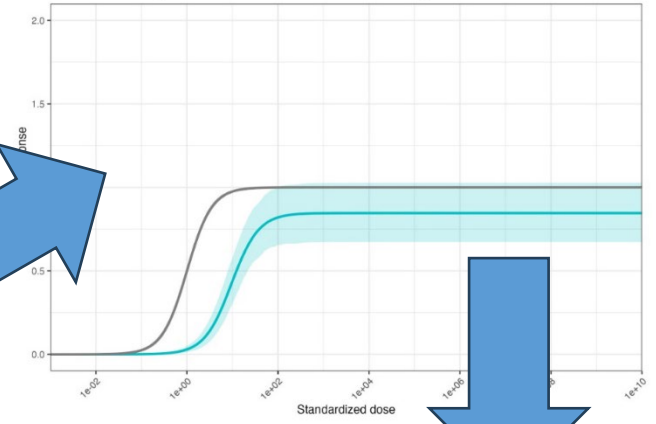
Bayesian Dose-Response Modeling and Standardization for all D/C-R Datasets



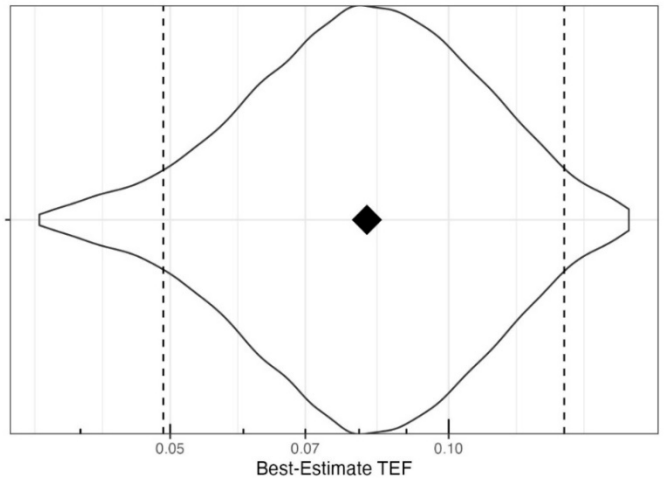
Synthesis of data using Bayesian Meta-Analysis



Derivation of Model Reference-Congener Relationship Estimates



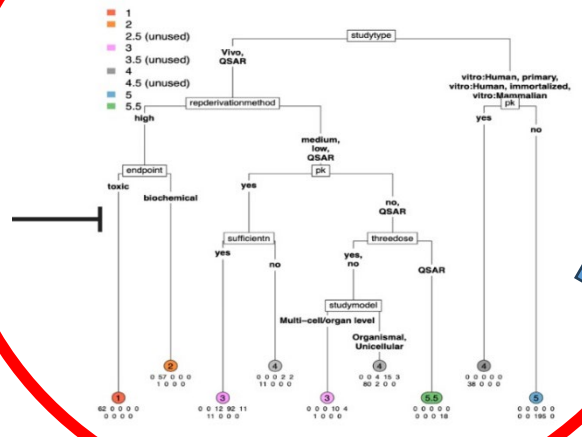
Derivation of Model TEF Estimates



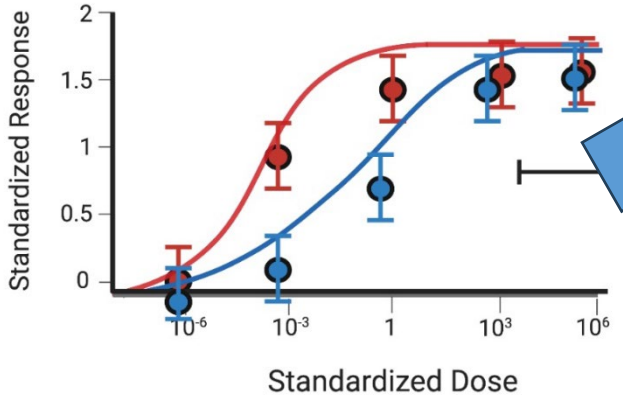
Ring et al. (2023)

Best Estimate TEF Workflow

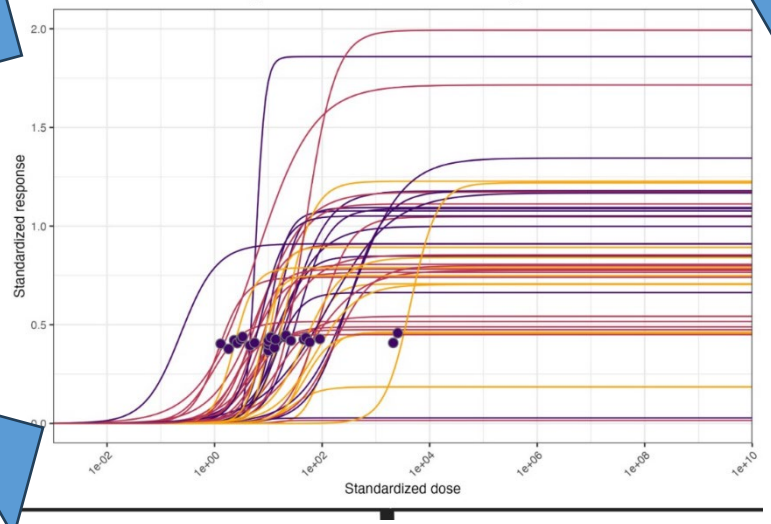
Machine Learning-Based REP Dataset Quality Predictions



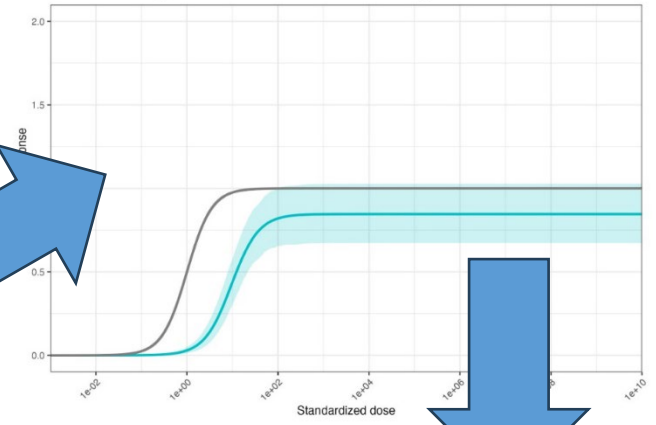
Bayesian Dose-Response Modeling and Standardization for all D/C-R Datasets



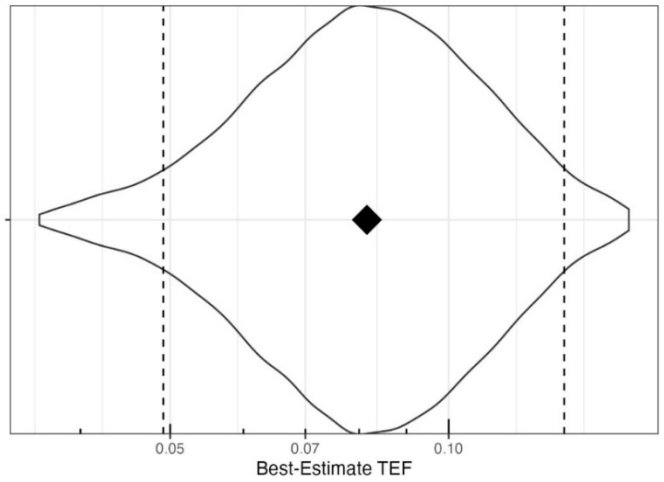
Synthesis of data using Bayesian Meta-Analysis



Derivation of Model Reference-Congener Relationship Estimates



Derivation of Model TEF Estimates



Ring et al. (2023)

EPA United States Environmental Protection Agency

Machine-Learning-Based REP Dataset Quality Predictions [Wikoff et al., 2023]

- Expert panel (2004): Identify study attributes that characterize reliability and relevance
- Expert panel (2004): Rate study quality on categorical scale from 1-5.5 (1 best)
  - based on qualitative expert judgment
  - no explicit decision criteria
- Train a machine-learning model to *infer* the expert panel's decision criteria & quantify uncertainty in category ratings
- (How to translate quality category into quantitative weight? That comes later!)

Study attributes of reliability/relevance [Wikoff et al., 2023]

- **Study type** (*in vivo,* or *in vitro* with human primary, human immortalized, or non-human mammalian cells)
- **Study endpoint** (toxic or biochemical)
- **Study model** (whole organism, organ-level, unicellular)
- Whether the congener had a **kinetic profile similar to TCDD**
- Whether the study duration was sufficient to achieve **kinetic steady-state**
- Whether a **sufficient number of dose levels** was tested (≥3)
- Whether a **sufficient number of animals/replicates** was tested (n depends on endpoint)
- Whether **maximal response** was achieved

Machine-learning model infers expert panel's decision criteria

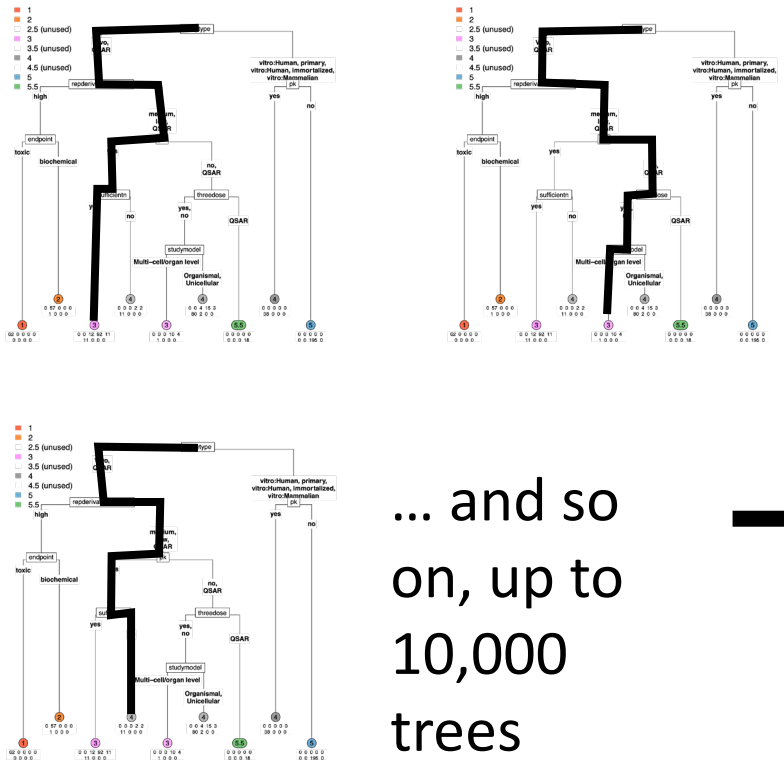
Random forest model
(ensemble of decision trees)

Features (Predictors)

Study attributes

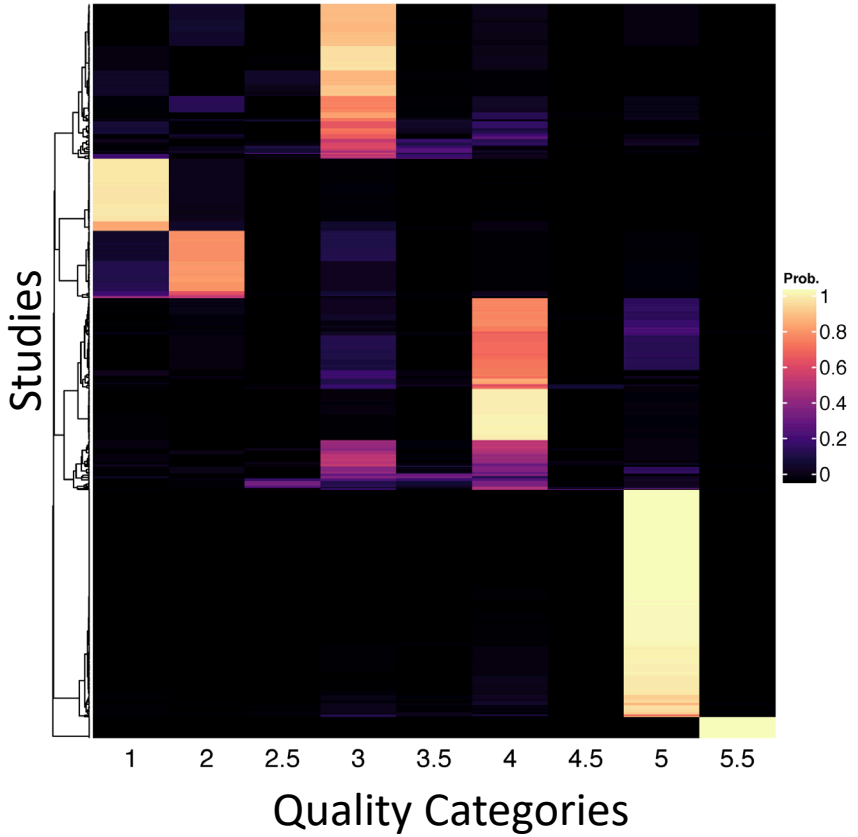
Target (Response)

Quality category 1-5.5 (assigned by expert panel)



… and so on, up to 10,000 trees

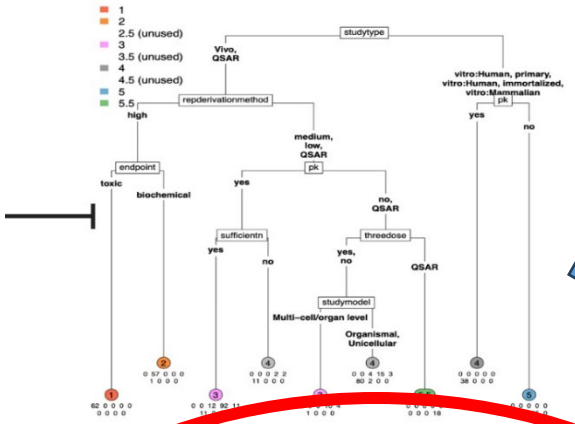
Result: "Probability" that each study is in each quality category



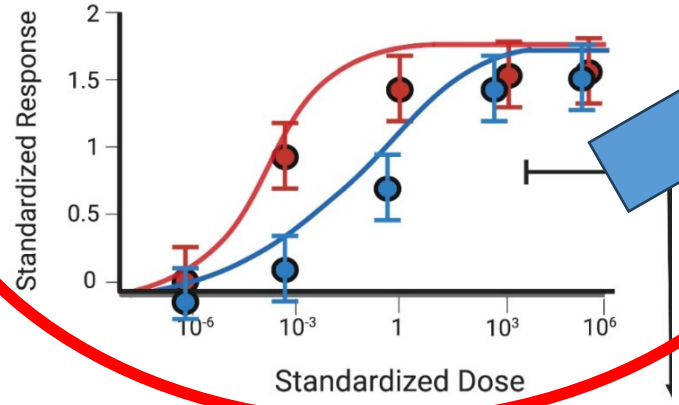
Wikoff et al., 2023

Best Estimate TEF Workflow

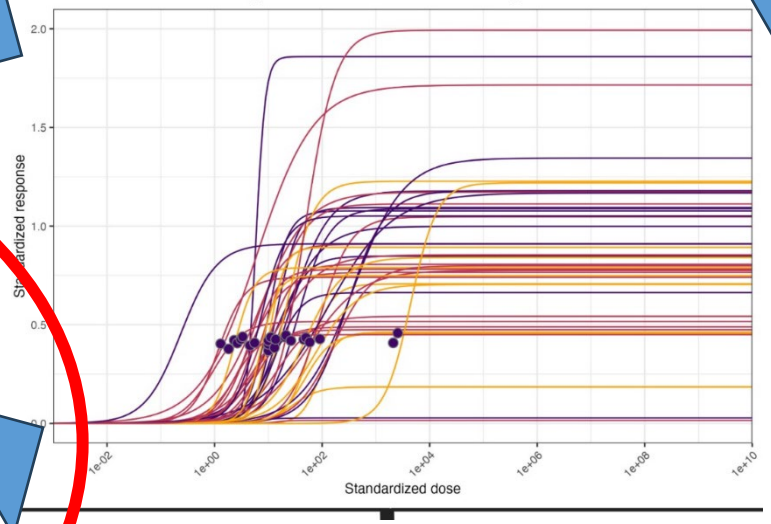
Machine Learning-Based REP Dataset Quality Predictions



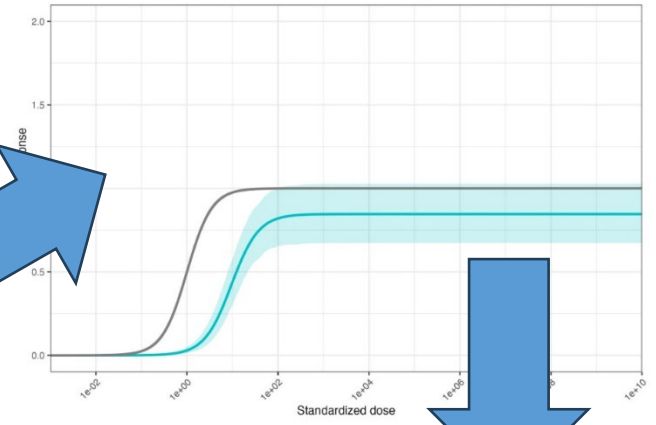
Bayesian Dose-Response Modeling and Standardization for all D/C-R Datasets



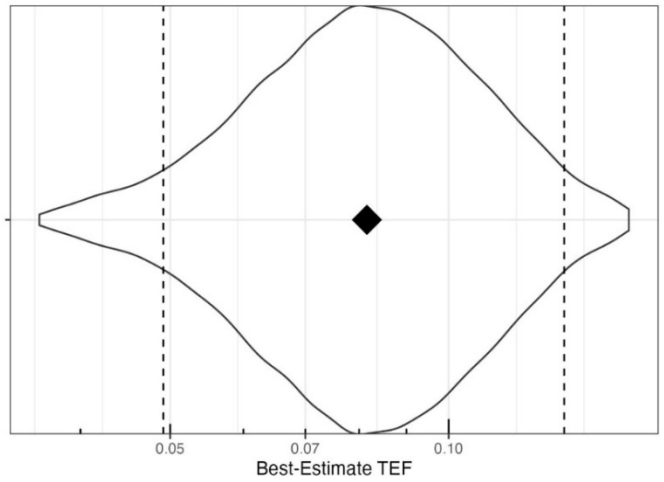
Synthesis of data using Bayesian Meta-Analysis



Derivation of Model Reference-Congener Relationship Estimates



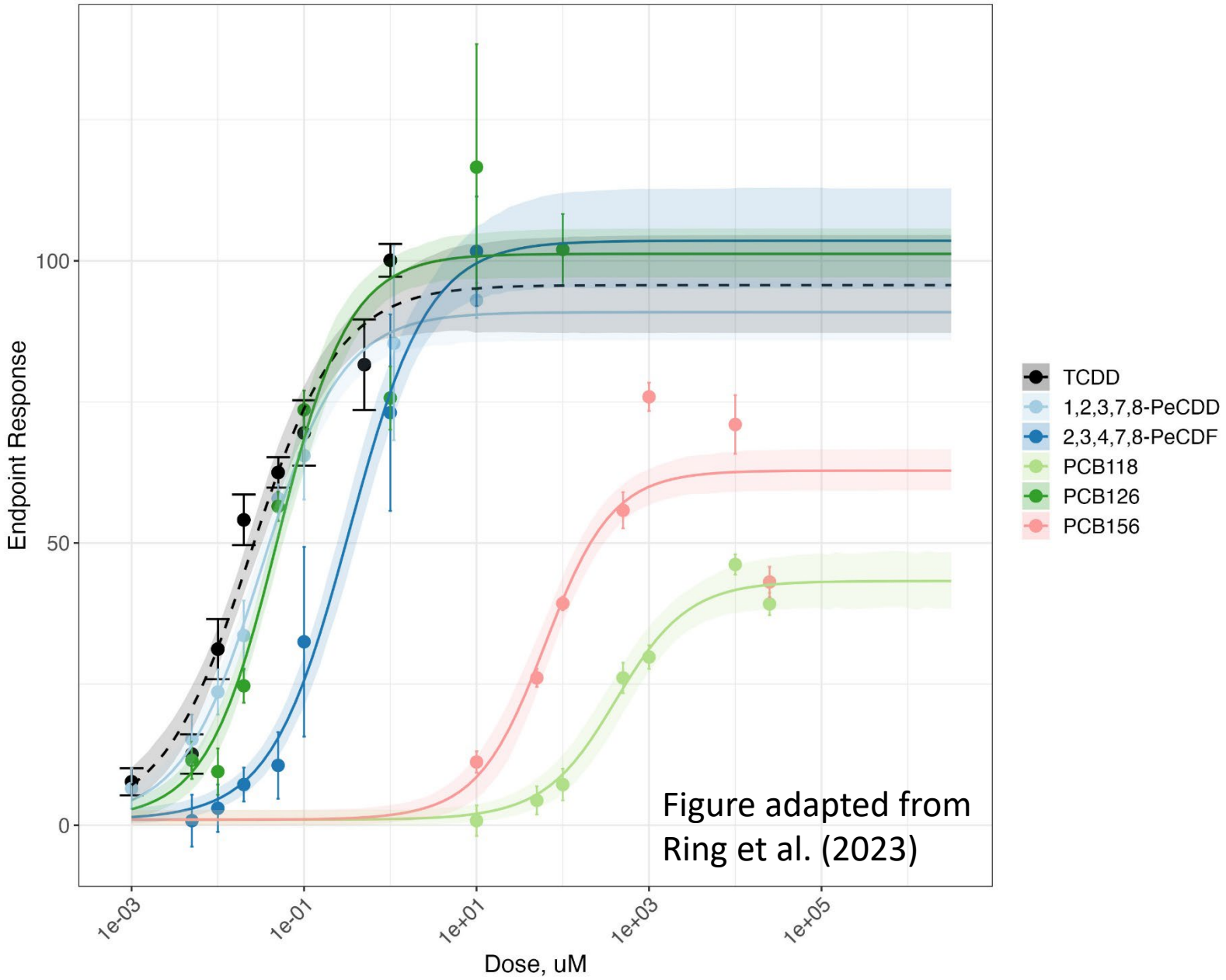
Derivation of Model TEF Estimates



Ring et al. (2023)

Bayesian Dose-Response Modeling

- Hill model
- Within each study, fit multiple congeners simultaneously
- Result: *Probabilistic* estimates of Hill model parameters, per study & congener

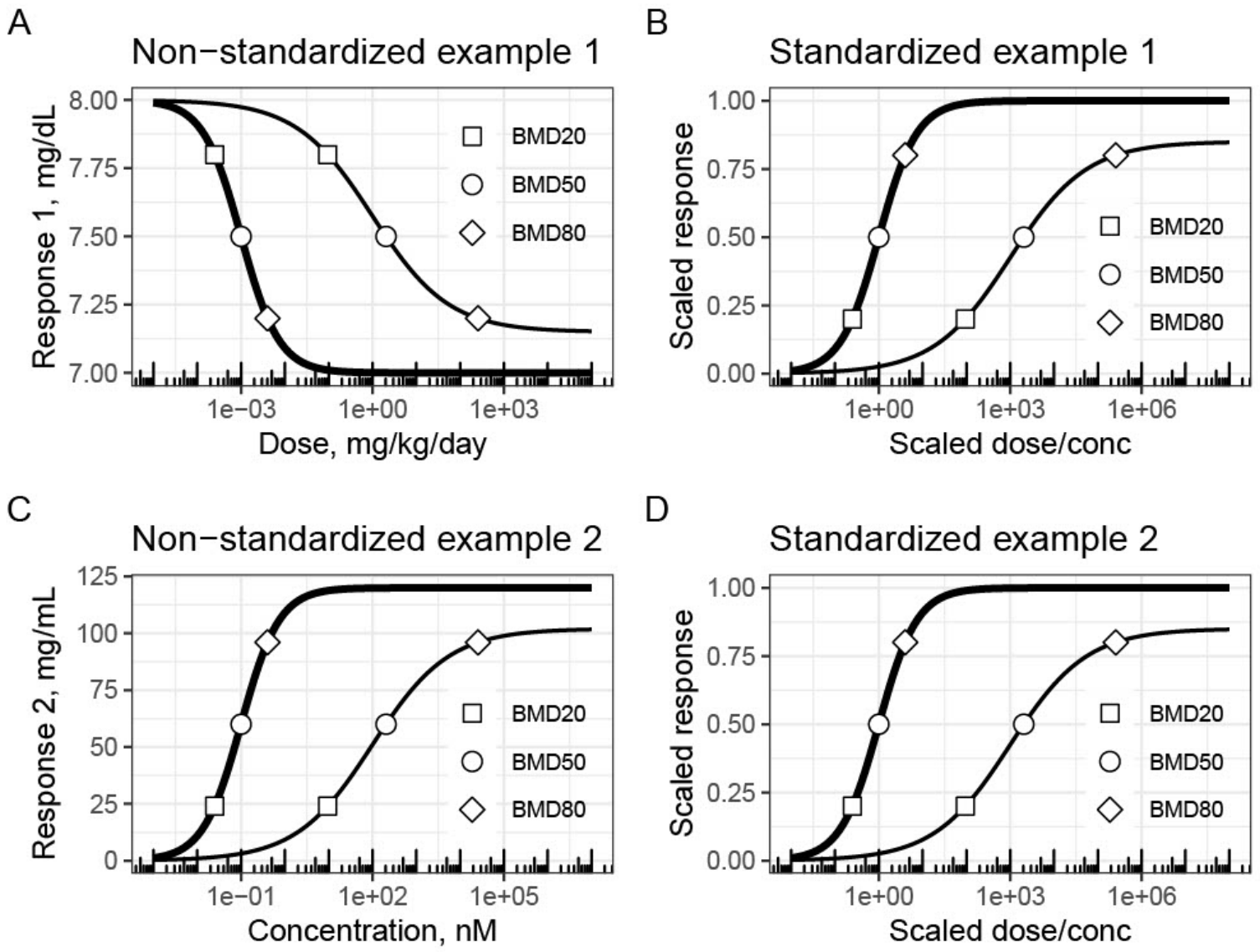


Standardize fitted dose-response curves

1. Subtract control response

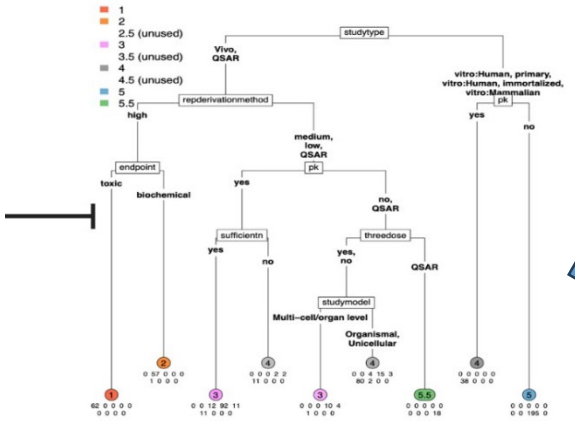
2. Normalize response to TCDD max

3. Normalize dose to TCDD ED50

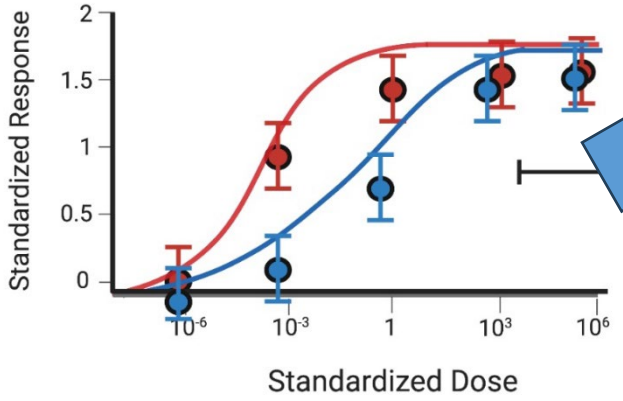


Best Estimate TEF Workflow

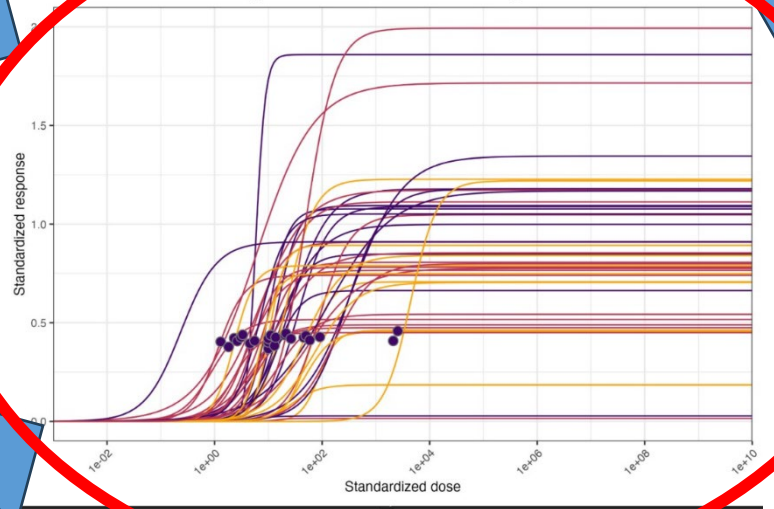
Machine Learning-Based REP Dataset Quality Predictions



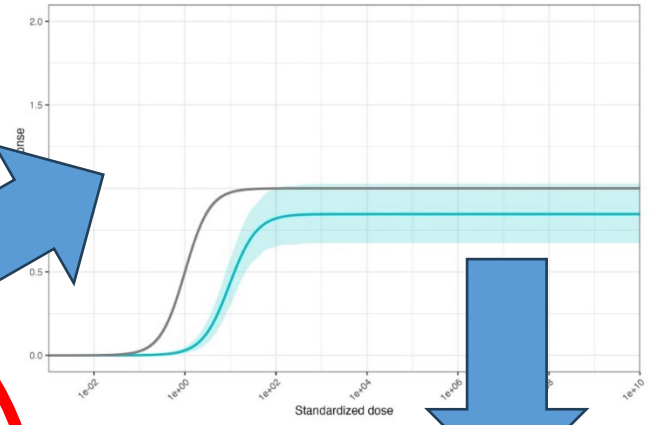
Bayesian Dose-Response Modeling and Standardization for all D/C-R Datasets



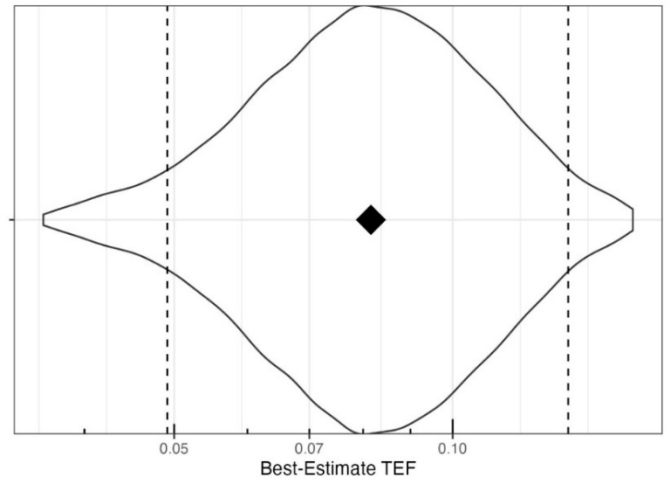
Synthesis of data using Bayesian Meta-Analysis



Derivation of Model Reference-Congener Relationship Estimates



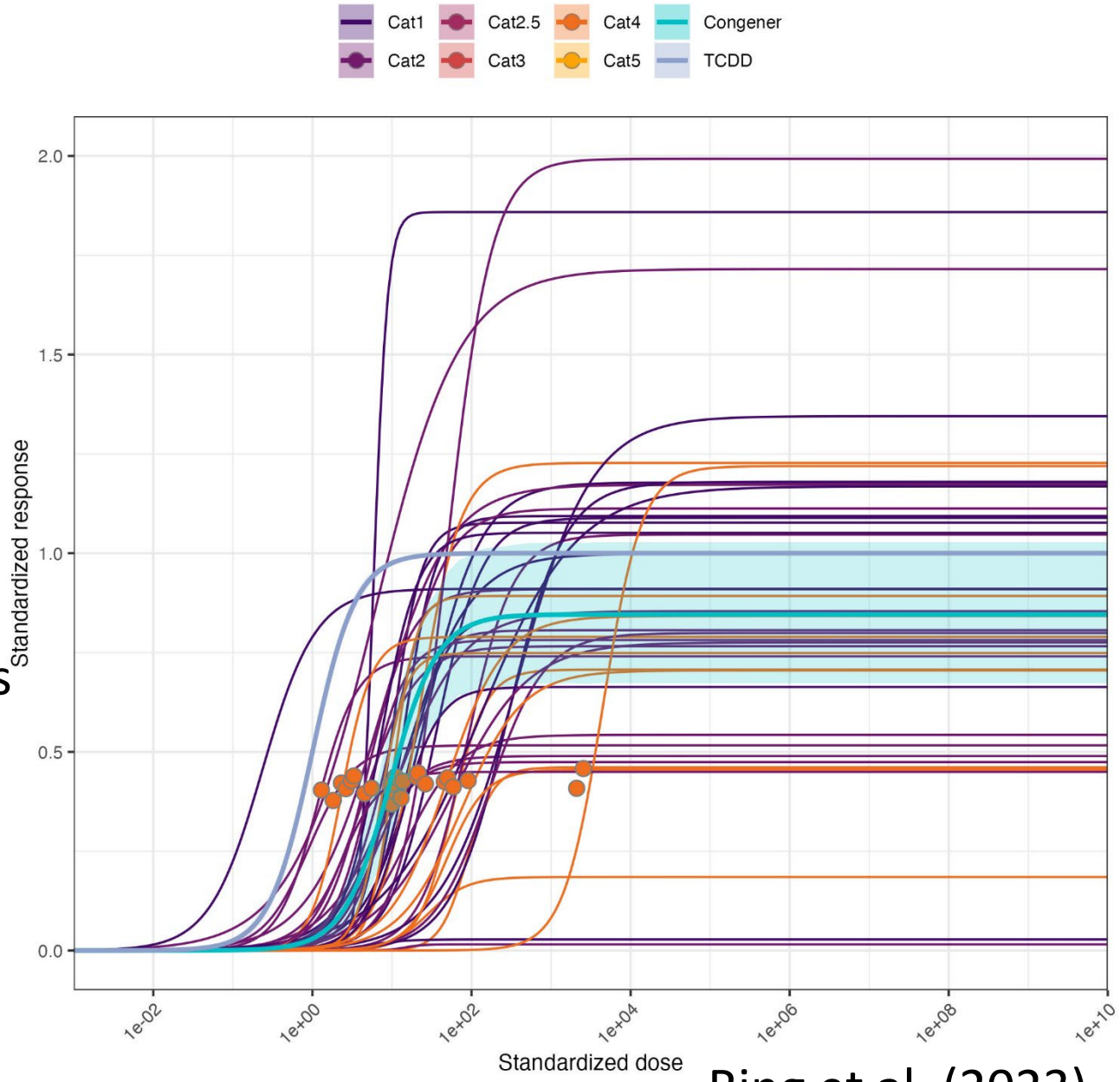
Derivation of Model TEF Estimates



Ring et al. (2023)

Synthesis of data using Bayesian Meta-Analysis

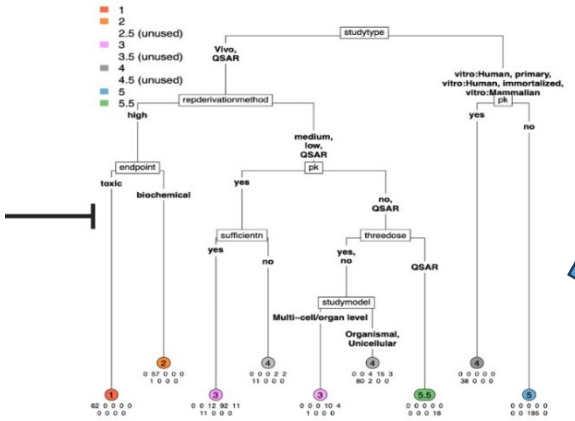
- Infer the “average” standardized dose-response curve for each congener (and its uncertainty) from all the study-specific curves
- **Quality weighting**: Assume higher-quality curves are clustered closer to “average” curve, lower-quality curves scattered more widely
- “Database uncertainty” represented by Bayesian priors: range of *possible* “average” curves assumed *a priori*



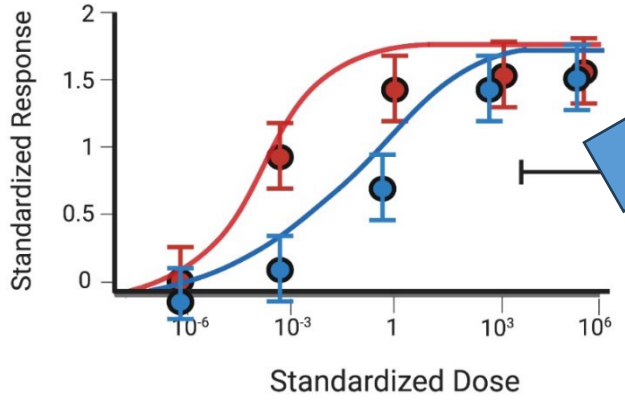
Ring et al. (2023)

Best Estimate TEF Workflow

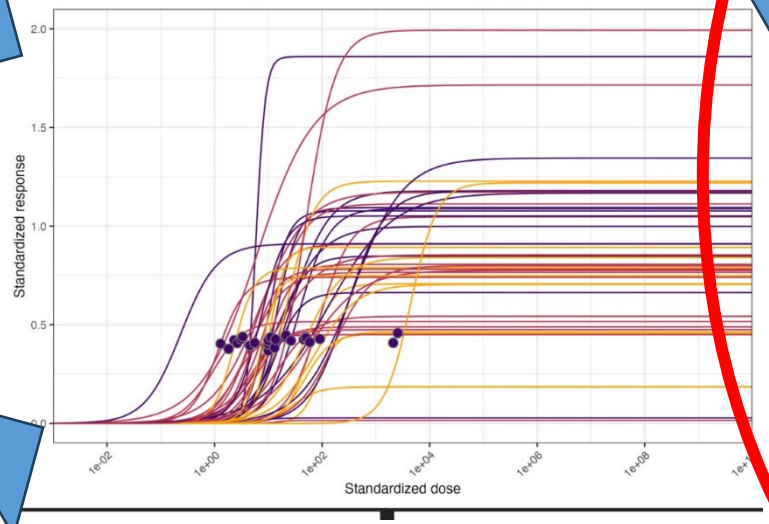
Machine Learning-Based REP Dataset Quality Predictions



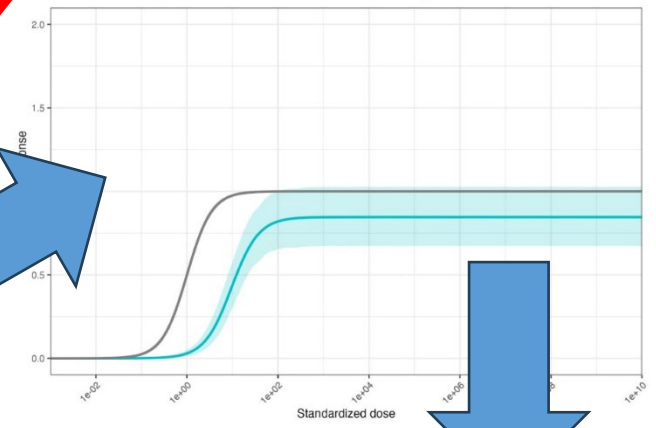
Bayesian Dose-Response Modeling and Standardization for all D/C-R Datasets



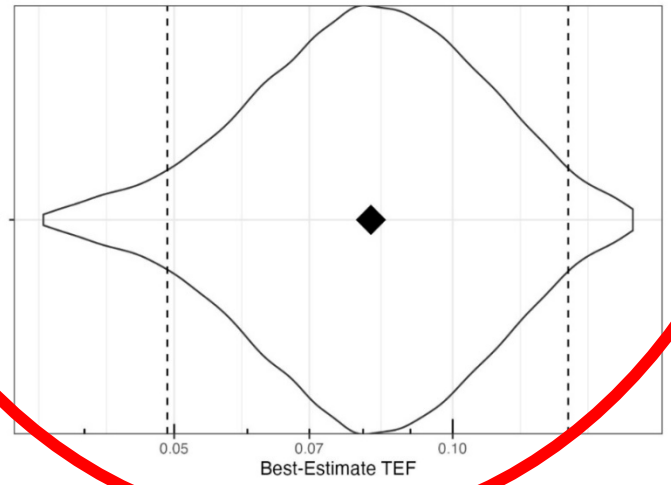
Synthesis of data using Bayesian Meta-Analysis



Derivation of Model Reference-Congener Relationship Estimates

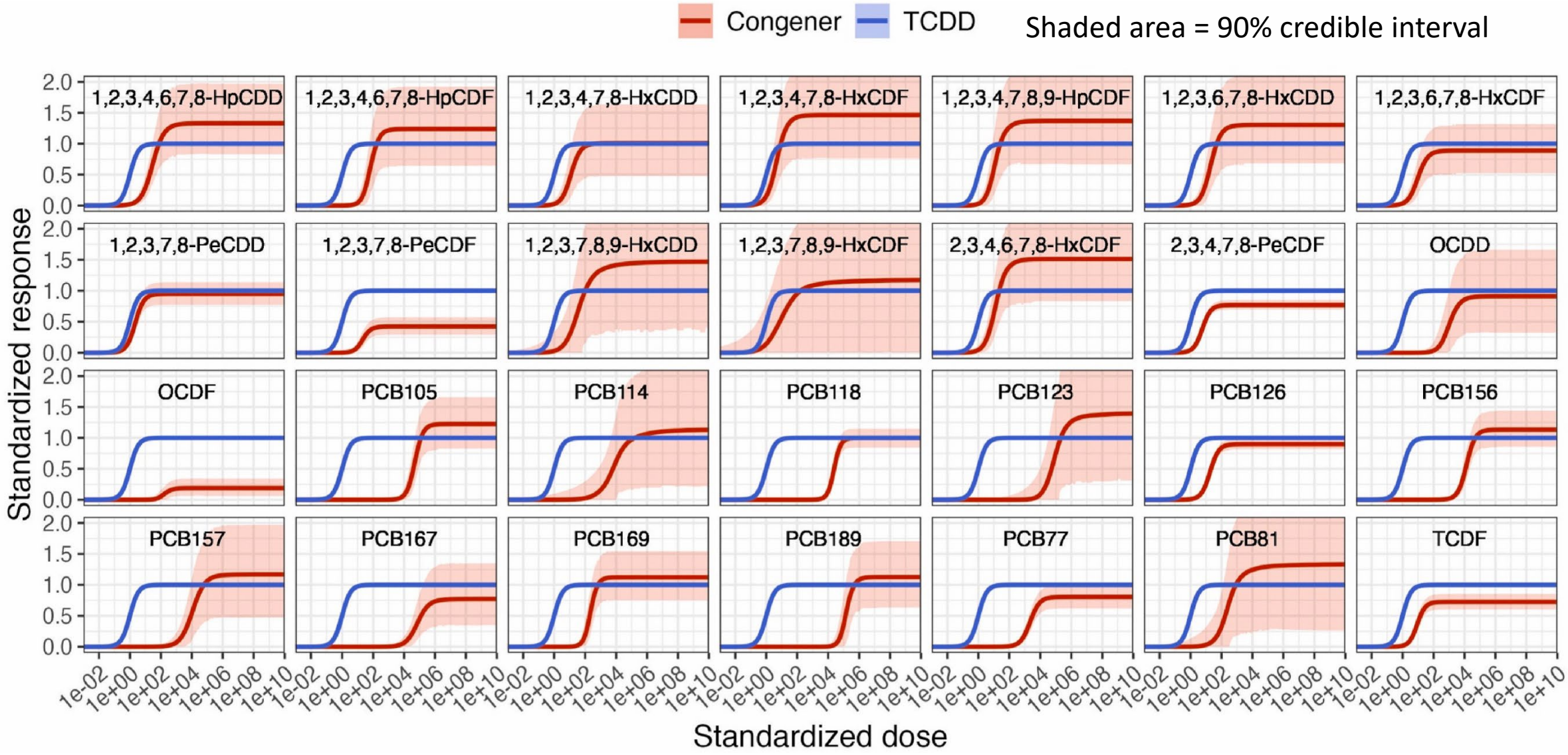


Derivation of Model TEF Estimates

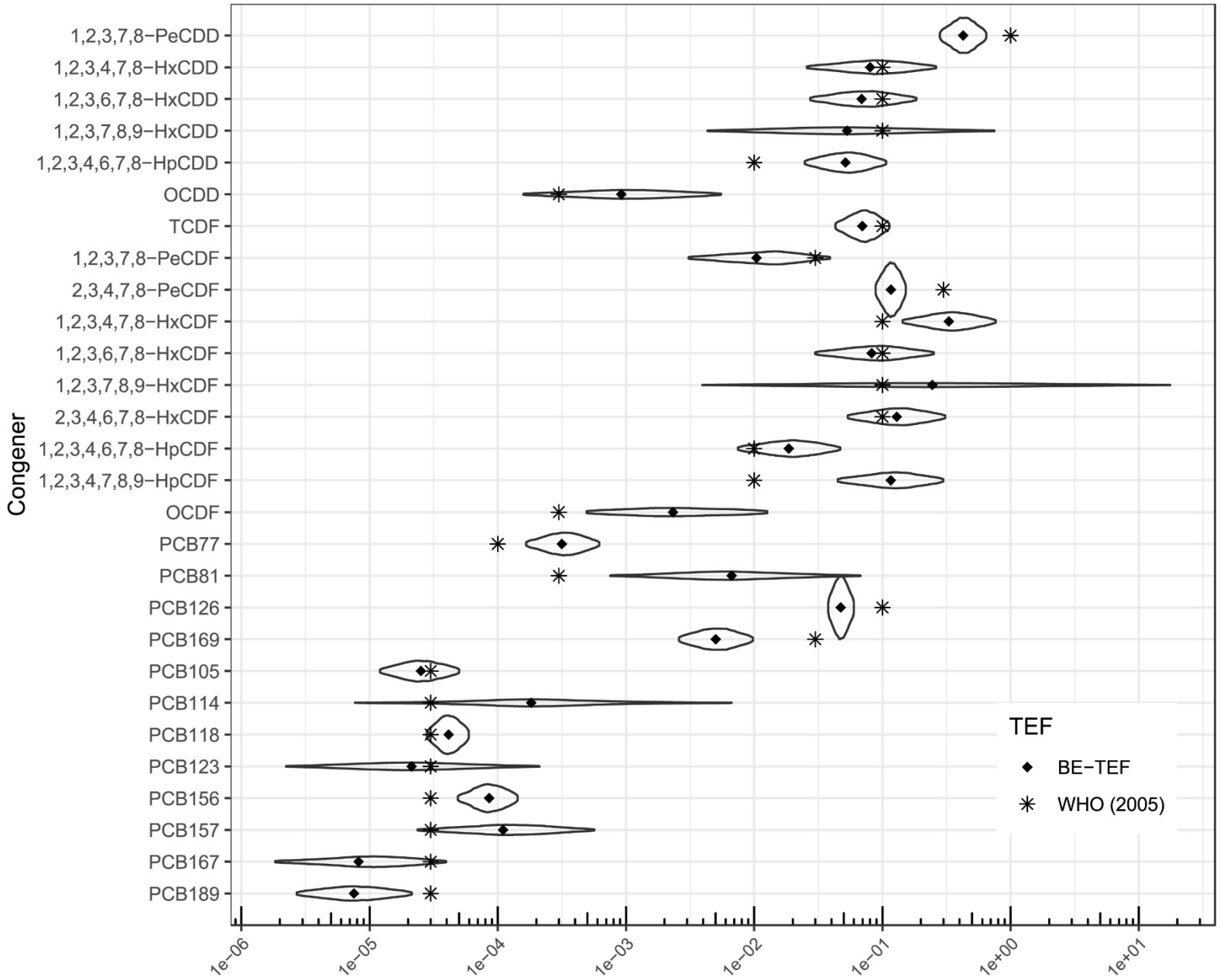


Ring et al. (2023)

Model Estimate of Standardized Dose-Response Relationship for each Congener (Fitch et al., 2023)



Best-Estimate TEFs and Uncertainty Distributions (Fitch et al. 2023)



October 2022: WHO expert panel re-evaluated TEFs for dioxin-like compounds

- Evaluated the Best-Estimate TEF workflow and the resulting TEF values
- WHO panel adopted “Best-Estimate” TEFs for everything except mono-ortho PCBs
- Outcome and details published in peer-reviewed article (DeVito et al., 2023)

Summary and Conclusion

- TEFs for dioxin-like compounds are estimated based on weight-of-evidence from a body of relative potency (REP) data
  - $REP_{2004} \rightarrow REP_{2021}$ (updated to include new REP studies!)
- REP studies are of varying reliability and relevance
- We developed a method to quantitatively integrate REP data
  - Consensus quantitative weighting by reliability & relevance
  - Integration of dose-response and non-dose-response REP data
- Best-Estimate TEF Workflow:
  - Transparent assumptions & model structure
    - Database & model code are proprietary, but described in published literature
  - Full quantification of uncertainty at every stage
- WHO (2022) expert panel agreed on applying this method in re-evaluating TEFs for dioxin-like compounds
- EPA is currently reviewing the WHO's recent reanalysis and update of the TEFs for dioxin and dioxin-like chemicals and determining their suitability for use in agency decision making

Disclosures

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.

This work was partially carried out in support of the 2022 WHO Expert Consultation on Updating the 2005 Toxic Equivalency Factors for Dioxin Like Compounds, Including Some Polychlorinated Biphenyls; methods and results presented herein were provided to the WHO expert working group for consideration during the 2022 evaluation.

ToxStrategies received funding from:

- Tierra Solutions, Inc., subsequently Glenn Springs Holdings, Inc. (2004-2018)
- European Food Safety Authority (2018-2021)

KeyToxicology received funding from:

- European Food Safety Authority

Some authors were also supported in part by:

- Intramural Research Program of the NIH (MDV, NW, LB, MW)
- National Cancer Institute (LB)
- U.S. Environmental Protection Agency (MDV, CR)

Dr. Ring was employed by ToxStrategies from June 2016-August 2020, and then became employed by the US EPA. After leaving ToxStrategies, she did not receive any form of payment from ToxStrategies or any of its clients for her continued work on this project.

Dr. Birnbaum is currently a defense expert in dioxin-related litigation.

The contents of this presentation reflect the opinions and views of the authors. The mention of trade names and commercial products does not constitute endorsement or use recommendations.

References

- Van den Berg, M. *et al.* (2006). The 2005 World Health Organization reevaluation of human and mammalian toxic equivalency factors for dioxins and dioxin-like compounds. *Toxicol Sci* **93**, 223-241, doi:10.1093/toxsci (2006).
- Haws, L. C. *et al.* (2006). Development of a refined database of mammalian relative potency estimates for dioxin-like compounds. *Toxicol Sci* **89**, 4-30, doi:10.1093/toxsci/kfi294 (2006).
- Raudenbush, S. W., & Bryk, A. S. (1985). Empirical Bayes meta-analysis. *Journal of Educational Statistics*, 10(2), 75–98.
- Wikoff, D. *et al.* (2023). Development and Application of a Systematic and Quantitative Weighting Framework to Evaluate the Quality and Relevance of Relative Potency Estimates for Dioxin-Like Compounds (DLCs) for Human Health Risk Assessment. *Regul Toxicol Pharmacol. 145,* 105500. https://doi.org/10.1016/j.yrtph.2023.105500
- Ring, C. *et al.* (2023). A Multi-tiered Hierarchical Bayesian Approach to Derive Toxic Equivalency Factors for Dioxin-Like Compounds. *Regul Toxicol Pharmacol 143,* 105464. https://doi.org/10.1016/j.yrtph.2023.105464
- Fitch, S. *et al.* (2023). Systematic Update to the Mammalian Relative Potency Estimate Database and Development of Best Estimate Toxic Equivalency Factors for Dioxin-like Compounds. *Regul Toxicol Pharmacol* . 147. https://doi.org/10.1016/j.yrtph.2024.105571
- DeVito, M., *et al.* (2024). The 2022 world health organization reevaluation of human and mammalian toxic equivalency factors for polychlorinated dioxins, dibenzofurans and biphenyls. *Regul Toxicol Pharmacol*, 146, 105525. https://doi.org/10.1016/j.yrtph.2023.105525
- U.S. Environmental Protection Agency. (2010). *Recommended Toxicity Equivalence Factors (TEFs) for Human Health Risk Assessments of 2,3,7,8Tetrachlorodibenzo-p-dioxin and Dioxin-Like Compounds. (EPA/100/R 10/005). Washington, D.C.: Risk Assessment Forum Retrieved from https://www.epa.gov/sites/default/files/2013-09/documents/tefs-for-dioxin-epa-00-r-10-005-final.pdf*