A Multi-tiered Hierarchical Bayesian New Approach Method to Derive Toxic Equivalency Factors for Dioxin-Like Compounds (Working Title)

**Supplemental Figures and Material**

Caroline Ring,1a Alexander Blanchette\*2; William D. Klaren2; Seneca Fitch3; Laurie Haws1; Matthew W. Wheeler4; Michael DeVito5, Nigel Walker4, Daniele Wikoff2

1. ToxStrategies, Austin, TX, USA

2. ToxStrategies, Asheville, NC, USA

3. ToxStrategies, Katy, TX, USA

4. National Institute of Environmental Health Sciences/National Institutes of Health, Research Triangle Park, NC, USA

5. Environmental Protection Agency, Center for Computational Toxicology and Exposure, Research Triangle Park, NC, USA

a. Current affiliation with Environmental Protection Agency, Center for Computational Toxicology and Exposure, Research Triangle Park, NC, USA

Corresponding Author:

Daniele Wikoff\*

dwikoff@toxstrategies.com
31 College Place, Suite B118

Asheville, NC 28801, USA

Phone: 828.348.6833

**Supplemental Figures**



**Supplemental Figure S1. REP Quality criteria as determined by the REP expert panel (citation) and implemented into the machine-learning model.** Adapted from Wikoff et al. (Submitted, 2023)



**Supplemental Figure S2. Derivation of Best-Estimate TEF Uncertainty distributions and point estimates from the predicted Best-Estimate D/C-R relationship between the reference and congener.** The Best-Estimate TEF point-estimate as derived from the ratio between the reference and congener curves at 50% response (left). The uncertainty in the horizontal position of the congener curve at 50% response that is imparted into the uncertainty surrounding the Best-Estimate TEF (right).

# Summary of Markov Chain Monte Carlo

Conceptually, Markov Chain Monte Carlo is a “guess and check” type of algorithm, which generates samples from the uncertainty distribution of the parameters as estimated from the data. First, the “guess”: the algorithm randomly draws a “proposed” set of parameter values. Then, the “check”: the algorithm assesses how well the proposed set of values fits the data, along with how well it fits any *a priori* assumptions (*e.g.*, assumptions that parameters must be within certain bounds). The algorithm then decides probabilistically whether to accept the proposed set of values, or to reject them and keep the previous set of values. If the proposed values are a good fit to both the data and the *a priori* assumptions, then they are more likely to be accepted (but not guaranteed acceptance). If the proposed values are a poor fit to the data, the *a priori* assumptions, or both, then they are less likely to be accepted (and may be summarily rejected if they violate a prior assumption, *e.g.* if a proposed value is out of bounds). If accepted, the proposed set of parameter values is retained as a sample from the uncertainty distribution for the parameters. If rejected, the proposed set of parameter values is discarded, and instead, the most recent *accepted* set of parameter values is *repeated* as a sample from the uncertainty distribution for the parameters.

Then, the algorithm randomly draws another set of proposed values, repeats the checking process, and repeats the probabilistic decision to accept or reject the new proposed values. This “guess and check” process is repeated many times, with each iteration producing one sample from the uncertainty distribution for the parameters as informed by the data (where some samples may be repeats of the previous sample, if the proposed values were rejected). The result is a large number of MCMC samples, constituting one Markov chain.

Each MCMC sample is one possible set of model parameters. Sets of values that are a better fit to the data and prior assumptions will appear more frequently; sets of values that are a poorer fit will appear less frequently (Brooks, 2011). The overall result is a large number of samples from the uncertainty distribution of the parameters as estimated from the data.

The goodness of fit to the data is checked using a *likelihood function*: a statistical model of the error between the model predictions and the observed data. The goodness of fit to the prior assumptions is checked using *priors*: statistical models that describe any *a priori* assumptions and uncertainty in the parameters.