



# FRAMEWORK FOR A QUANTITATIVE SYSTEMIC TOXICITY MODEL

Lisa Truong<sup>1</sup>, Gladys Ouedraogo<sup>2</sup>, Sophie Loisel-Joubert<sup>2</sup>, Matt Martin<sup>1</sup>

Lisa Truong <sup>1</sup> [Truong.Lisa@epa.gov](mailto:Truong.Lisa@epa.gov) | 919.541.3560

<sup>1</sup> US EPA- ORD – NCCT , RTP, NC 27711 | <sup>2</sup> L'Oréal, Paris, France

## BACKGROUND

- US EPA is collaborating with L'Oréal to develop high throughput screening (HTS) and non-animal testing methods to assess potential systemic toxicity of chemical compounds
- EPA ToxCast project is analyzing data generated on >1,000 chemicals across rapid, automated HTS assays with human gene and protein targets
- Toxicity Reference Database (ToxRefDB) is a repository of >5,000 legacy animal studies on ~1,000 chemicals, and captures the animal studies using a standardized, multilayered effect vocabulary across various study types and species

## OBJECTIVE

Utilize the ToxCast HTS *in vitro* and ToxRefDB *in vivo* data to develop a quantitative model predictive of systemic toxic effects

## METHODS

### CHEMICAL SELECTION

- Chemicals from ToxRefDB were filtered to include only systemic endpoints and study type and species adjustment parameters. Lowest effect levels (LELs) were obtained and utilized in the modeling.

### Systemic Endpoint Parameters

Effect Category: Parental, Systemic and Maternal

Route of Administration: Oral

Species: Mouse, Rat, Primate, Dog, Rabbit, Hamster

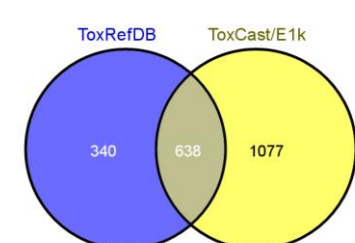


Figure 1. Venn Diagram Visualizing Chemicals with Data in ToxRefDB and ToxCast. 638 chemicals were used in model development since both data sources available. 145 chemicals were removed for future external validation.

**Orders of Magnitude Uncertainty (OMU) Bounds:** OMU = 4 \* standard deviation

### ANALYSIS

All analysis and figures were produced using R (<http://www.r-project.org/>) with packages (ggplot2, reshape2)

#### Chemical Profile

The lowest LEL for each chemical was obtained from ToxRefDB and expressed in log<sub>10</sub> form using the formula:

$$-\log_{10} \frac{\text{adjusted LEL}}{10000}$$

#### Read Across

Toxprint fingerprints were obtained from <https://toxprint.org/> and applied to the 978 chemicals. Pearson's correlation was used to compute similarity between chemicals for the 729 fragments. The nearest 5 neighbors were utilized to compute a mean nearest neighbor LEL (NNL).

#### ToxCast Assays and Reverse Toxicokinetics

913 assays were assigned into 72 biological groupings. Each assay belonged only to one group. Assignment of assays were based on their Pearson's correlation coefficient to other assays. An average activity level was computed per biological grouping. Steady state concentration (CSS) of humanized rat serum was utilized to calculate an oral equivalence dose per biological grouping for chemicals with CSSS values available.

#### Parameter Selections

Univariate analysis either Pearson's correlation coefficient or linear regression was applied to identify predictive variables for modeling

## QUANTITATIVE MODEL FRAMEWORK

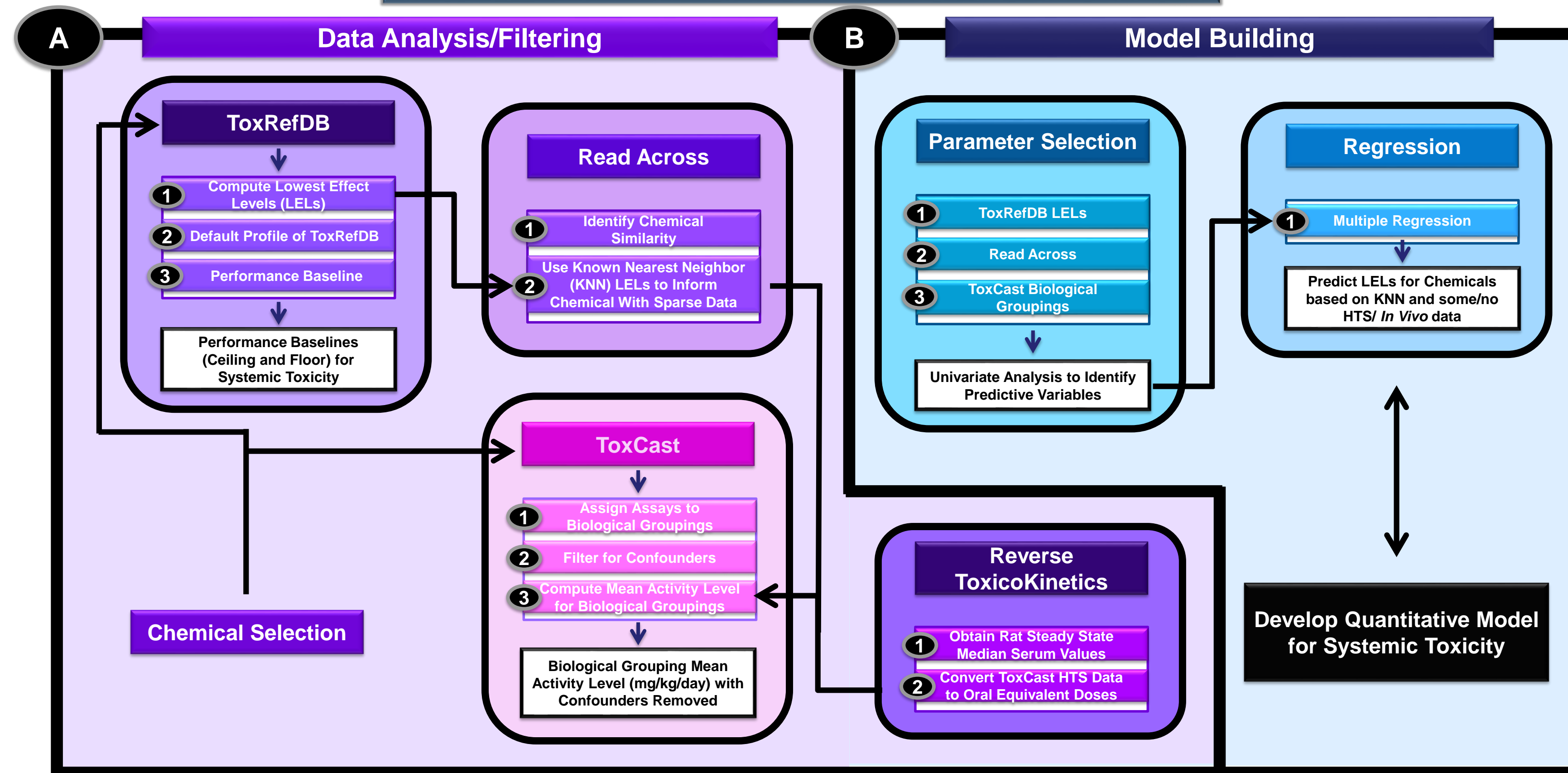


Figure 2. Framework to Developing a Quantitative Model for Systemic Toxicity.

The first portion of the framework (A) is to select chemicals from two sources: ToxRefDB, and ToxCast. Within ToxRefDB, the analysis conducted yields a performance baseline for systemic toxicity, while for ToxCast, the outcome will be to have biological groupings with a mean activity level in mg/kg/day with confounders removed (white boxes). Read across and reverse toxicokinetics are applied to aid in the process. After data analysis (B), the model is built by using univariate analysis to identify predictive values from the three data sources. The parameters selected are applied to a linear regression to predict LELs for chemicals.

## DEFAULT BASELINE OF TOXREFDB

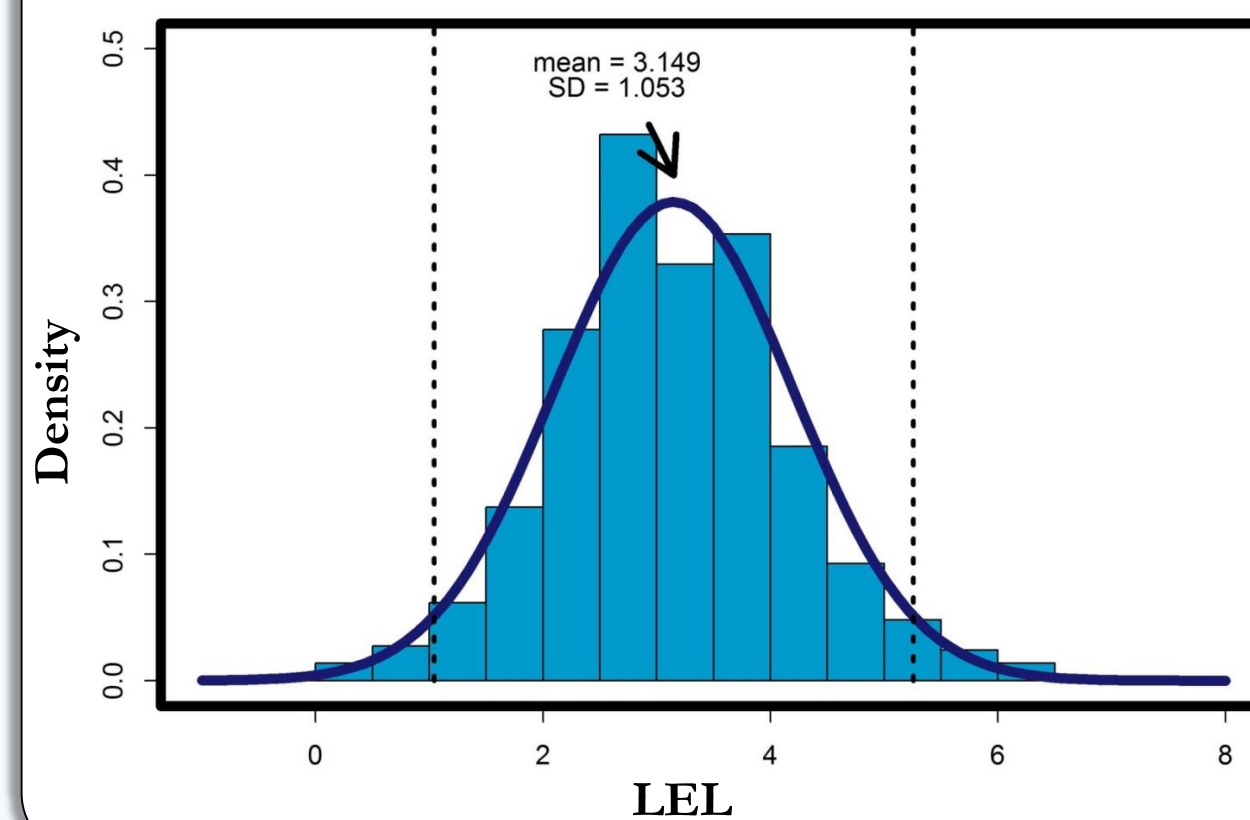


Figure 3. Profile of 603 Chemicals with Systemic Toxicity LELs and ToxCast Data.

Of the 638 chemicals with available data from both sources, 603 fit the systemic filtration parameter set for ToxRefDB. The mean on a  $-\log_{10}$  scale for the 603 chemicals is 3.149 with a standard deviation (SD) of 1.053 and an OMU of 4.212. The dotted line represents mean  $\pm$  2 SD.

## TOXCAST BIOLOGICAL GROUPINGS

Table 1. ToxCast Biological Groupings Utilized in the Model

Biological Groupings	# Assay	Biological Groupings	# Assay	Biological Groupings	# Assay
androgen_receptor	5	estrogen_receptor_agonist	7	nuclear_receptor_non_steroidal	26
androgen_receptor_agonist	4	estrogen_receptor_antagonist	3	nuclear_receptor_non_steroidal_agonist	11
androgen_receptor_antagonist	2	estrogen_receptor_binding	3	nuclear_receptor_non_steroidal_antagonist	3
androgen_receptor_binding	3	glucocorticoid_receptor	8	nuclear_receptor_steroidal_agonist	3
aryl_hydrocarbon	2	gpcr_inhibition_other	2	oxidative_stress	15
basic_leucine_zipper_activation	10	gpcr_inhibition_other_receptor	6	p53_activation	5
cell_adhesion_inhibition	2	gpcr_inhibition_rhodopsin_like_receptor	68	p53_inhibition	3
cell_adhesion_molecule_activation	19	immune_cytokine_activation	21	peroxisome_proliferator_activated_receptor_alpha	3
cell_adhesion_molecule_inhibition	17	immune_cytokine_inhibition	17	peroxisome_proliferator_activated_receptor_gamma	7
cell_morphology	29	ion_channel_inhibition	7	phosphatase_activation	19
constitutive_androstane_receptor	5	kinase_activation	56	phosphatase_inhibition	19
constitutive_androstane_receptor_agonist	1	kinase_inhibition	60	pregnane_x_receptor	9
cyp_activation	30	ligand_gated_ion_channel_inhibition	13	proliferation	5
cyp_inhibition	31	matrix_metalloproteinase_activation	6	protease_activation	14
cytokine_activation	30	matrix_metalloproteinase_inhibition	9	protease_inhibition	18
cytokine_inhibition	40	misc_enzyme_activation_binding	4	receptor_tyrosine_kinase_activation	1
cytotoxicity	47	misc_enzyme_inhibition_binding	3	retinoic_acid_receptor	8
cytotoxicity_antagonist	6	misc_protein_activation	16	retinoic_acid_receptor_agonist	1
deacetylase_activation	5	misc_protein_inhibition	11	retinoic_acid_receptor_antagonist	1
deacetylase_inhibition	5	monamine_oxidase_activation	4	synthase_activation	1
dna_binding	26	monamine_oxidase_inhibition	4	synthase_inhibition	1
esterase_activation	6	neurotransmitter_transporter_inhibition	6	transferase_activation	1
esterase_inhibition	6	no_biological_pathway	43	transferase_inhibition	1
estrogen_receptor	10	non_basic_leucine_zipper_binding	14	transporter_inhibition	5

## READ ACROSS APPROACH

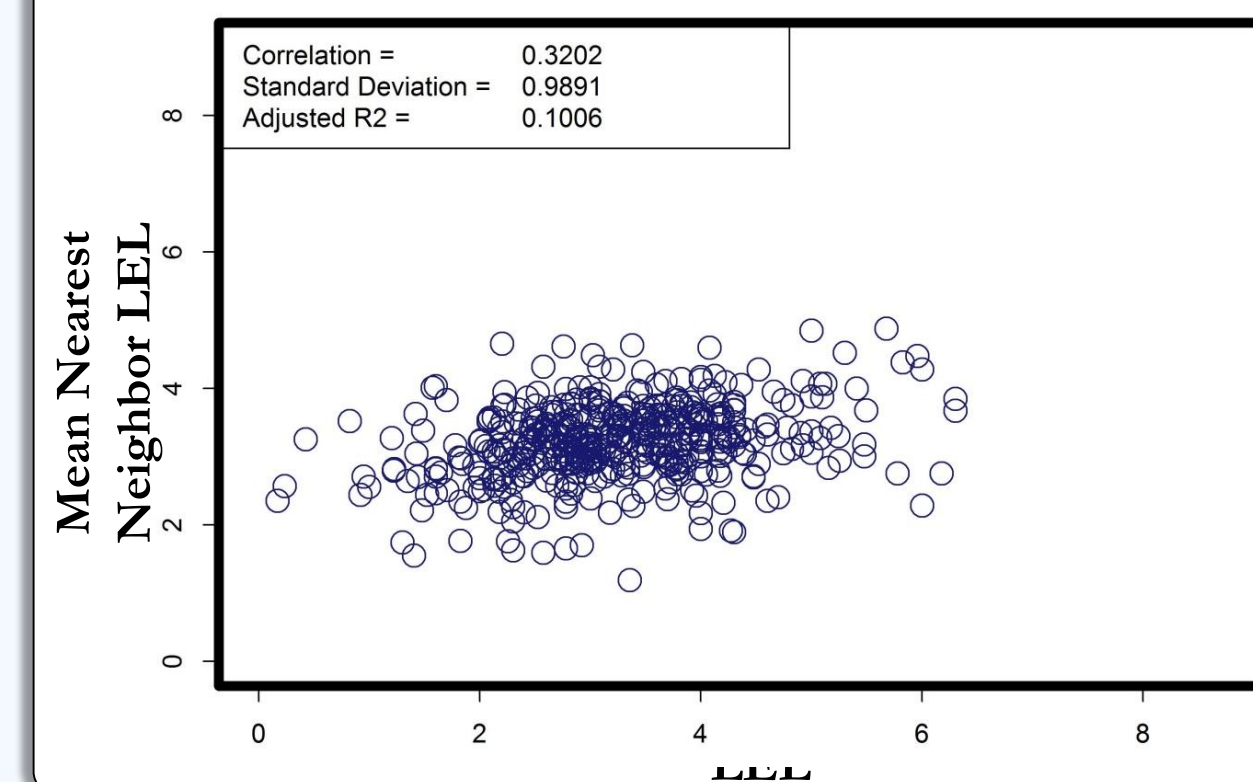


Figure 4. Read Across Correlation between Known Nearest Neighbor LEL Using ToxPrint Fingerprints and *In Vivo* LELs.

ToxPrint fingerprints were applied to the 603 chemicals, and the LELs of 5 nearest neighbors were used to compute a mean nearest neighbor LEL (NNL). The correlation coefficient for this approach is 0.3202 with a standard deviation of 0.9891 and an adjusted R2 of 0.1006. The OMU is 3.95.

## PRELIMINARY QUANTITATIVE MODEL

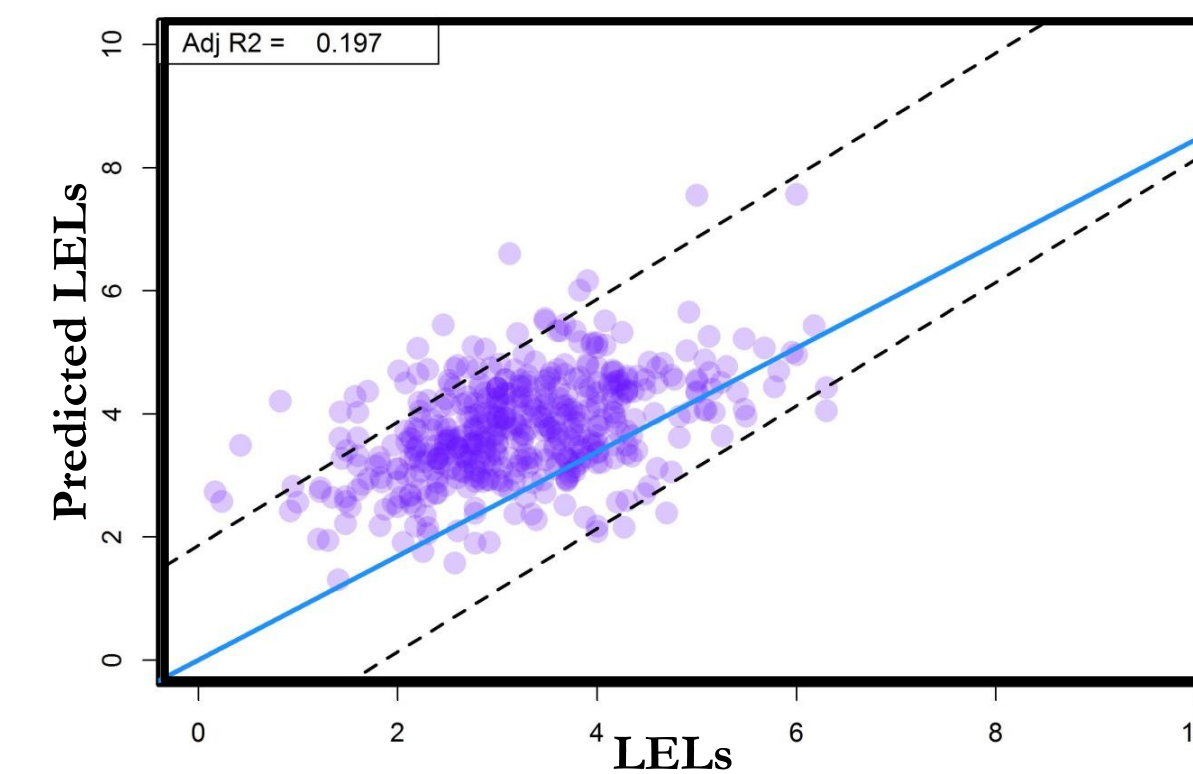


Figure 5. Regression Model Predicting LELs.

Using predictive variables selected by univariate analysis using LELs, read across and ToxCast biological groupings, a regression model (no RtK) was developed for the 603 chemicals. The adjusted R2 is 0.197, with residuals of 0.9 resulting in a OMU of 3.6

## MODEL PERFORMANCE EVALUATION

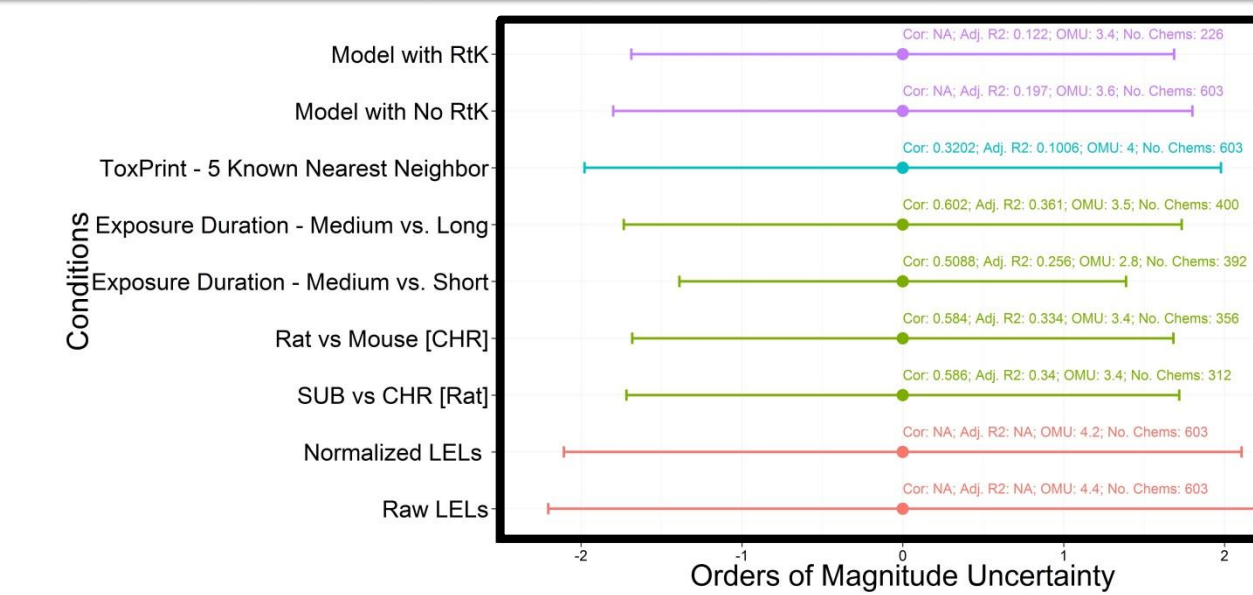


Figure 6. Model Performance Plot Comparisons to Floor and Ceiling Baseline.

A visualization of the orders of magnitude uncertainty (OMU) for 4 parameters (Default ToxRefDB, Performance Baseline, Read Across and Model; 4 colors) and the condition/comparisons conducted for that parameter.

## CONCLUSIONS/ FUTURE DIRECTIONS

- Incorporation of ToxRefDB, ToxCast, and Read Across data provides signal and does reduces the OMU from just utilizing *in vivo* data.
- With the framework established, future efforts will be focused on integrating reverse toxicokinetics for the full data set (603) chemicals and application of different read across methodologies.

