Environ Health Perspect

DOI: 10.1289/EHP2998

Supplemental Material

Conditional Toxicity Value (CTV) Predictor: An *In Silico* Approach for Generating Quantitative Risk Estimates for Chemicals

Jessica A. Wignall, Eugene Muratov, Alexander Sedykh, Kathryn Z. Guyton, Alexander Tropsha, Ivan Rusyn, and Weihsueh A. Chiu

Table of Contents

Figure S4. Example of mechanistic interpretation of QSAR model for the RfC and IUR. For each toxicity value, the top panel ranks the molecular descriptors by their frequency of use in the model. The top twenty are denoted by the dashed lines, and are shown separately in the middle panel with the descriptor names. The bottom panel compares the descriptor values for the top two descriptors between the highest and lowest potency toxicity values. RfC = Reference Concentration; IUR = Inhalation unit risk; QSAR = quantitative structure activity relationship. Definitions of each molecular descriptor can be found online at: https://cdk.github.io/cdk/1.5/docs/api/org/openscience/cdk/qsar/descriptors/molecular/package-summary.html

Figure S5. Comparison of Conditional Toxicity Value (CTV)-based (A, C, E) or High throughput screening (HTS) assay-based (B, D, F) risk characterization with "gold standard" risk characterization based on regulatory toxicity values. In each panel, the x-axis is the margin of exposure (MOE=Toxicity Value/Exposure) or Hazard Quotient (HQ=Exposure / Toxicity Value) derived from CTV (left panels) or HTS assays (right panels), which is compared to the MOE or HQ derived using regulatory toxicity values on the y-axis. Comparisons are made for regulatory NOAELs (panels A and B), BMDLs (panels C and D), or RfDs (panels E and F). In all cases, the predictions from CTV are based on cross-validation (panels A, C, and E). Each panel also includes lines indicating equality and 10-fold greater or less than equality (grey solid and dotted lines), nominal risk characterization thresholds (MOE = 100; HQ = 1), the number of compounds *n*, and the adjusted R^2 based on a linear model of log-transformed toxicity values. RfD = Reference Dose; NOAEL = No observed adverse effect level; BMDL = Benchmark dose lower confidence limit; OED05 = High throughput screening-based oral equivalent dose lower 5% confidence limit.

Additional File- Excel Document

References

Supplementary Methods and Results

Illustrative risk characterizations

In order to explore the possible risk assessment implications of using the Conditional Toxicity Value (CTV) predictor as compared to other methods, we calculated illustrative risk characterization values using (1) CTV predictions, (2) high throughput screening (HTS) assay-based oral equivalent dose (OED) estimates from Wetmore (2015), and (3) the "gold standard" regulatory NOAEL, BMDL, or RfD values. Risk characterization values require exposure estimates, so for illustration, we used the upper 95% exposure estimate from ExpoCast as the exposure value (Sipes et al. 2017; Wambaugh et al. 2013). We then calculated margins of exposure (MOEs) between that level of exposure and the NOAEL or BMDL (for CTV and "gold standard" regulatory values) and between exposure and the 5^{th} percentile OED_{05} (for HTS). We also calculated hazard quotients (HQs) as the ratio between exposure and the RfD for CTV and "gold standard" regulatory values. For HQs based on HTS assay-based results, we used a nominal "uncertainty factor" of 1000 for illustration, so that the HTS-based "RfD" = $OED_{05}/1000$. This value is based on the idea of (Crump et al. 2010) that RfDs based on in vitro studies could be derived by applying an additional uncertainty factor for in vitro-to-in vivo extrapolation. We then evaluated the degree to which CTV- and HTS-based risk characterizations replicated the risk values calculated using the "gold standard" regulatory toxicity values. This evaluation was related both to the consistency with "gold standard" regulatory values, as well as whether they gave different "decision" outcomes based on whether they satisfied the criteria of MOE > 100 or HQ < 1.

The results of these risk characterization illustrations are shown in Supplemental Figure S5. In all cases, as with the original toxicity values described in the main text, the CTV predictions for NOAELs (n=36) and BMDLs (n=14) resulted in MOEs that were more accurate and more precise (smaller absolute deviations and larger R^2) than MOEs based on HTS assays and IVIVE, when

compared to “gold standard” POD-derived MOEs. Risk characterizations using the RfD involve calculating a hazard quotient (HQ) instead of a MOE, and were available for more compounds (n=51), with similar results. Interestingly, for none of the compounds did the risk characterization using the “gold standard” regulatory toxicity values indicate a concern, defined by MOE < 100 or HQ > 1. These results were also the case for the risk characterizations based on CTV-derived toxicity values. On the other hand, HTS-based risk characterizations flagged some compounds as having a risk concern, suggesting that such risk characterizations may be more “conservative.”

Overall, when compared to the “gold standard” of using regulatory toxicity values, CTV gives more precise and more accurate risk characterizations than those derived from HTS assays and IVIVE. HTS-based risk characterizations tended to be more “conservative,” in that some compounds were flagged as having a potential risk whereas both the “gold standard”- or “CTV”-derived risk characterizations indicated acceptable MOEs or HQs. However, these results should be considered illustrative, given the additional assumptions and uncertainties involved in these calculations (e.g., exposure values, minimum MOE, uncertainty factor for HTS-based RfDs) as compared to the direct comparison of predicted toxicity values described in the main text.

Supplementary Figures

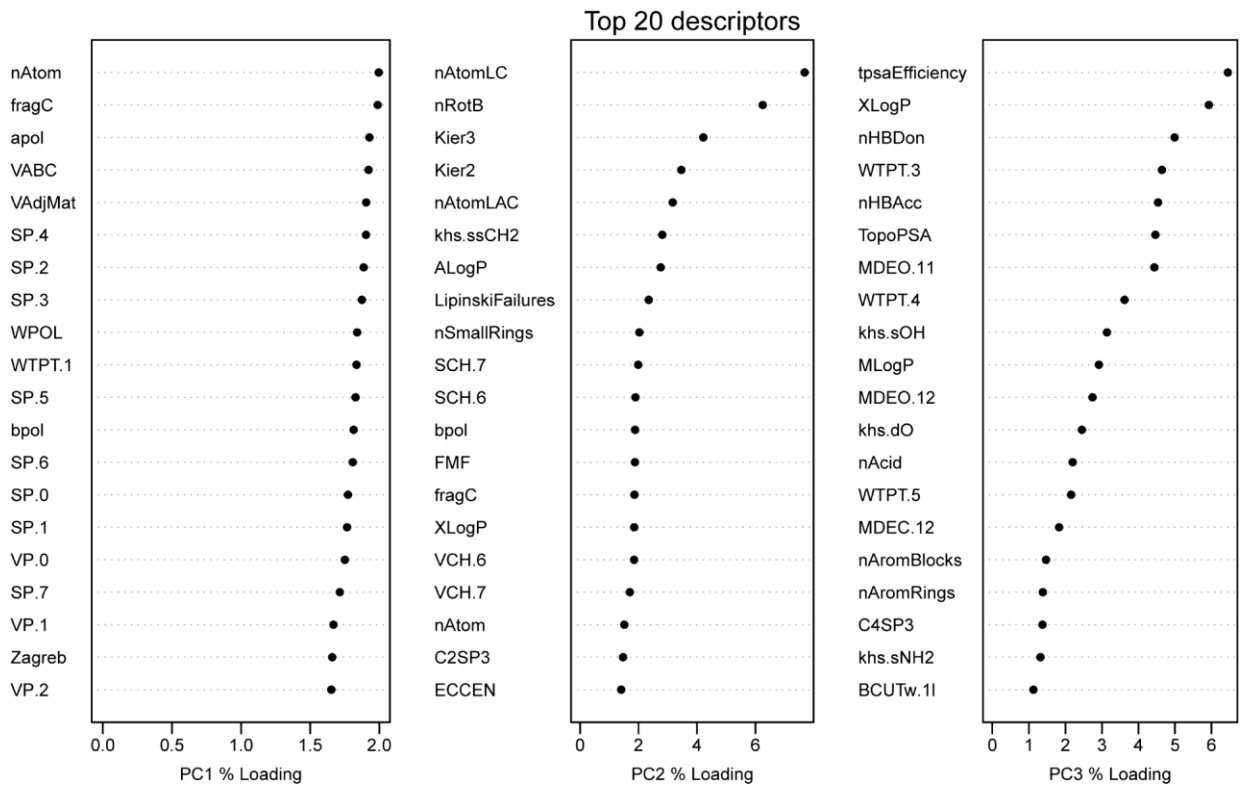

Figure S1. Principal component analysis loadings. The top twenty descriptors in each of the first three principal components are shown, with their percentage contributions. Definitions of each molecular descriptor can be found online at: https://cdk.github.io/cdk/1.5/docs/api/org/openscience/cdk/qsar/descriptors/molecular/package-summary.html

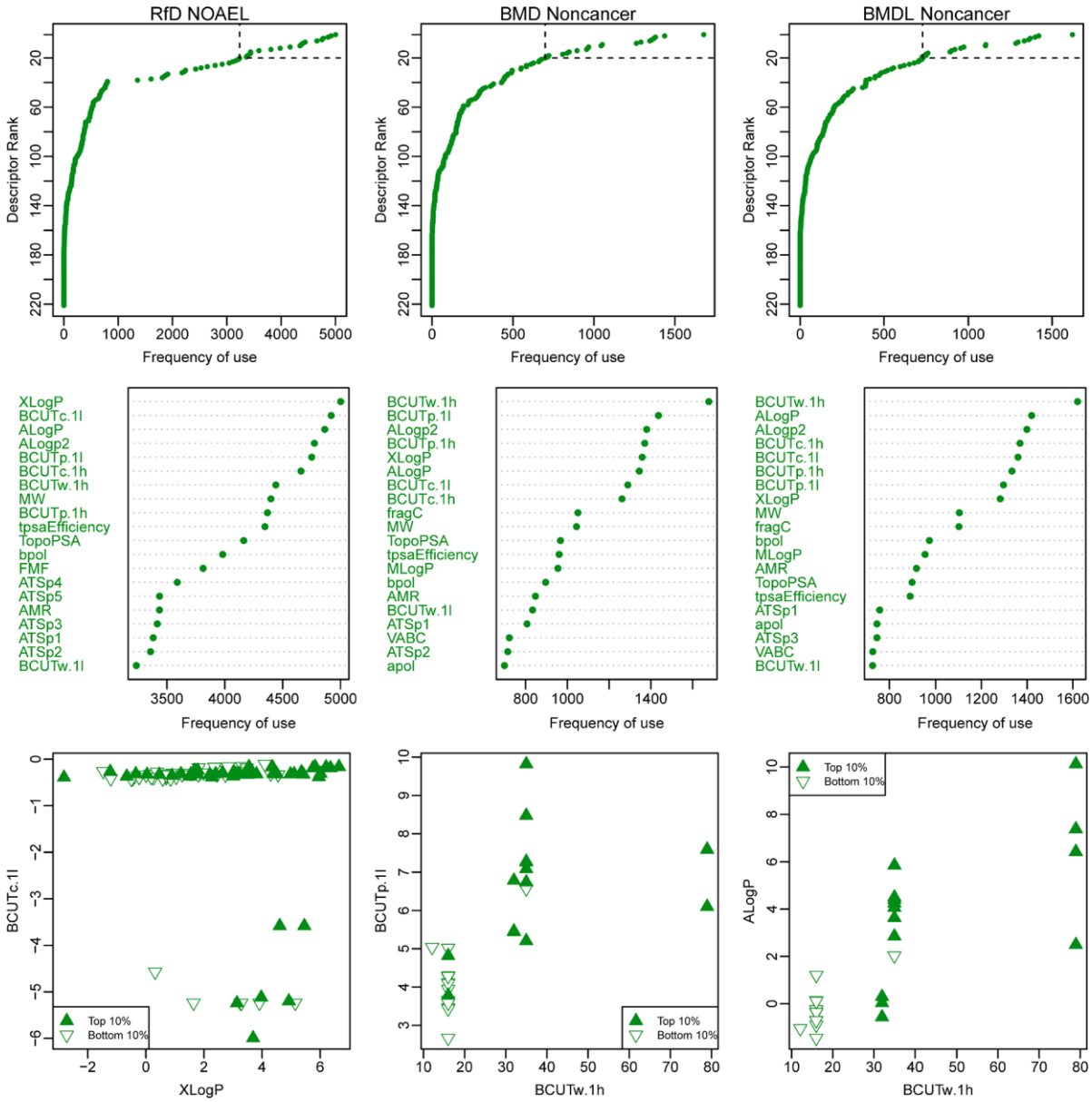

Figure S2. Example of mechanistic interpretation of QSAR model for the RfD NOAEL, BMD, and BMDL. For each toxicity value, the top panel ranks the molecular descriptors by their frequency of use in the model. The top twenty are denoted by the dashed lines, and are shown separately in the middle panel with the descriptor names. The bottom panel compares the descriptor values for the top two descriptors between the highest and lowest potency toxicity values. RfD = Reference Dose; NOAEL = No observed adverse effect level; BMD = Benchmark dose; BMDL = Benchmark dose lower confidence limit; QSAR = quantitative structure activity relationship. Definitions of each molecular descriptor can be found online at: https://cdk.github.io/cdk/1.5/docs/api/org/openscience/cdk/qsar/descriptors/molecular/package-summary.html

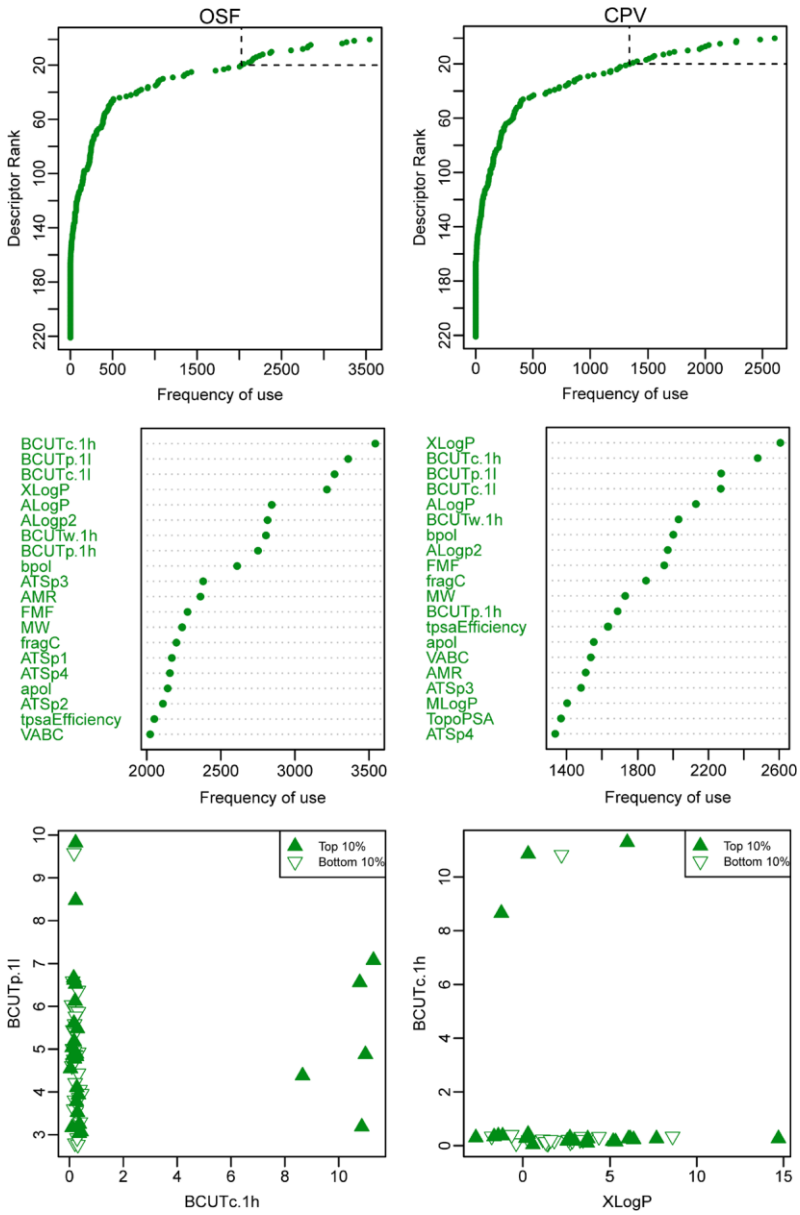

Figure S3. Example of mechanistic interpretation of QSAR model for the OSF and CPV. For each toxicity value, the top panel ranks the molecular descriptors by their frequency of use in the model. The top twenty are denoted by the dashed lines, and are shown separately in the middle panel with the descriptor names. The bottom panel compares the descriptor values for the top two descriptors between the highest and lowest potency toxicity values. OSF = Oral slope factor; CPV = Cancer potency value; QSAR = quantitative structure activity relationship. Definitions of each molecular descriptor can be found online at: https://cdk.github.io/cdk/1.5/docs/api/org/openscience/cdk/qsar/descriptors/molecular/package-summary.html

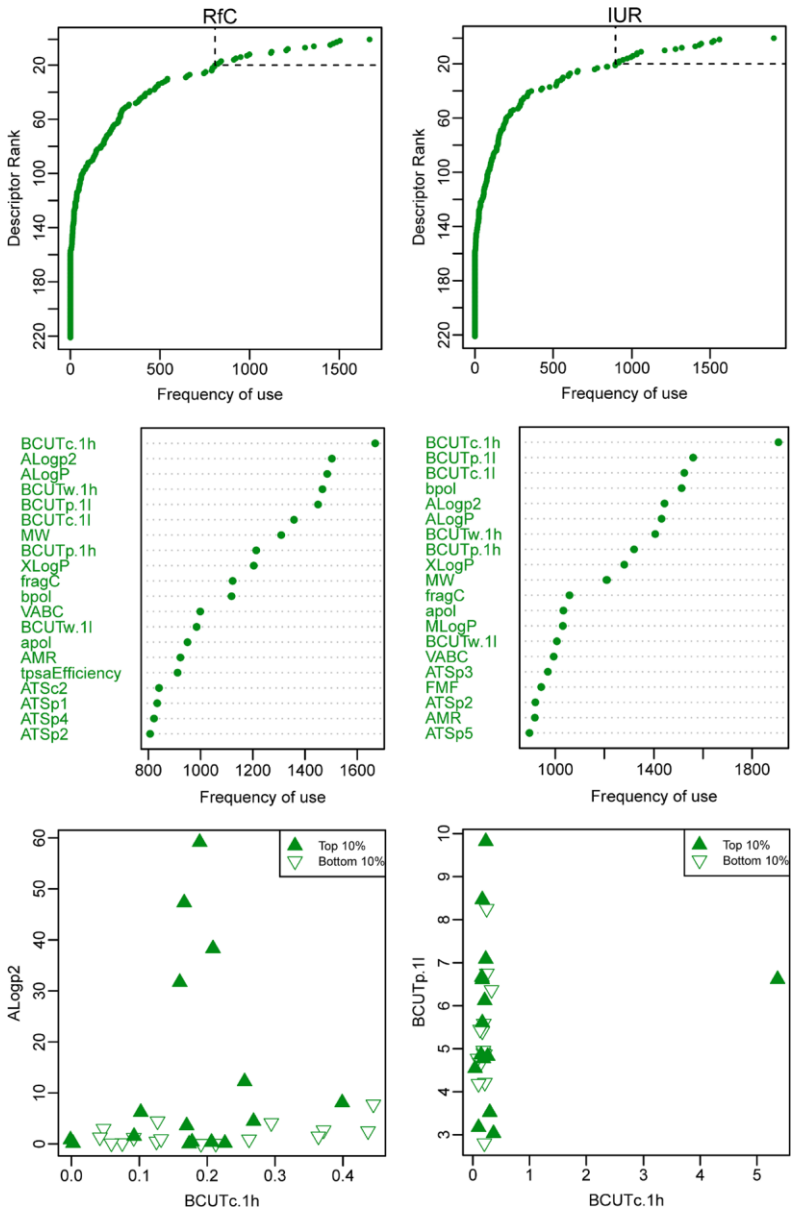

Figure S4. Example of mechanistic interpretation of QSAR model for the RfC and IUR. For each toxicity value, the top panel ranks the molecular descriptors by their frequency of use in the model. The top twenty are denoted by the dashed lines, and are shown separately in the middle panel with the descriptor names. The bottom panel compares the descriptor values for the top two descriptors between the highest and lowest potency toxicity values. RfC = Reference Concentration; IUR = Inhalation unit risk; QSAR = quantitative structure activity relationship. Definitions of each molecular descriptor can be found online at: https://cdk.github.io/cdk/1.5/docs/api/org/openscience/cdk/qsar/descriptors/molecular/package-summary.html

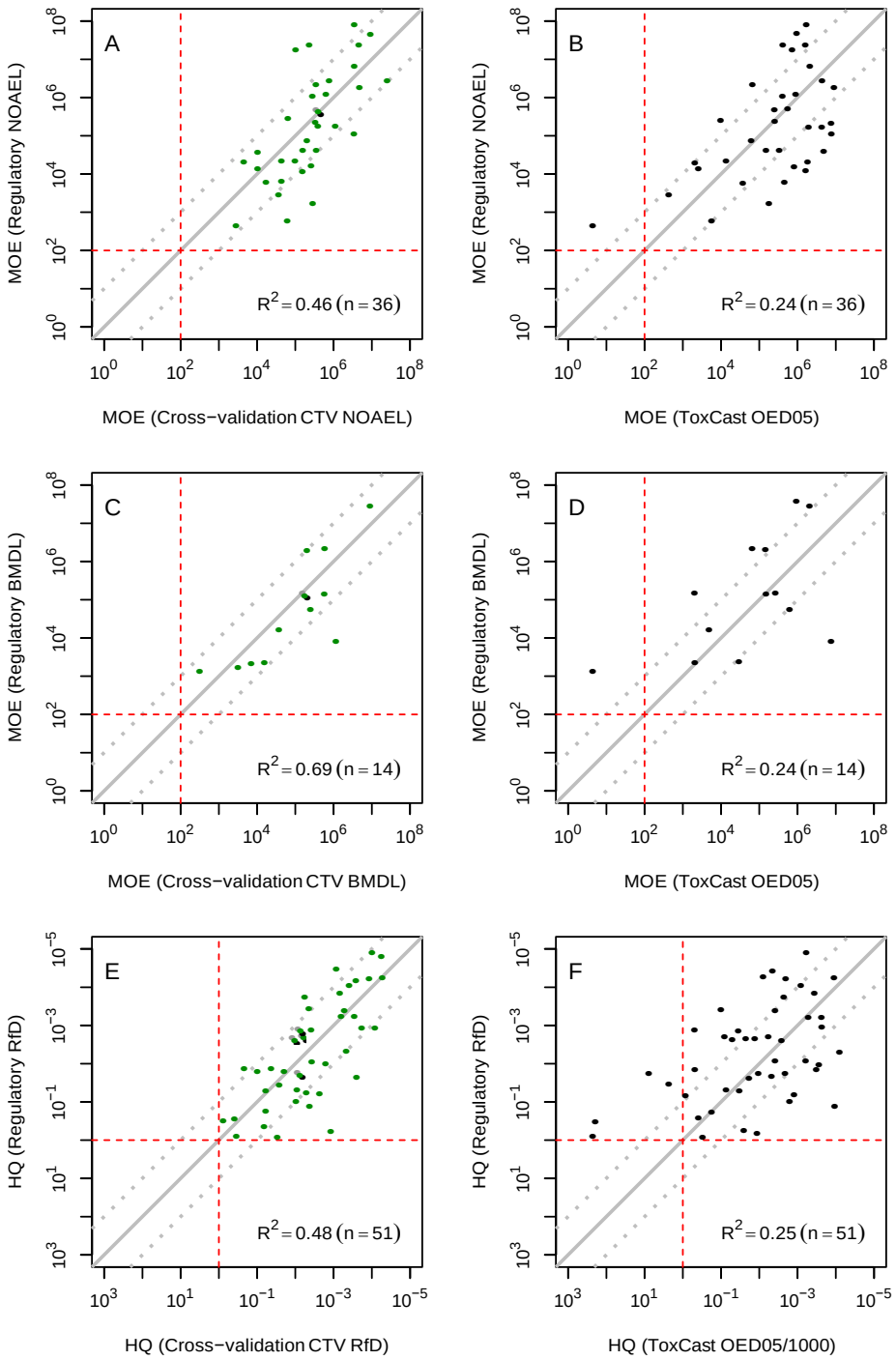

Figure S5. Comparison of Conditional Toxicity Value (CTV)-based (A, C, E) or High throughput screening (HTS) assay-based (B, D, F) risk characterization with "gold standard" risk characterization based on regulatory toxicity values. In each panel, the x-axis is the margin of exposure (MOE=Toxicity Value/Exposure) or Hazard Quotient (HQ=Exposure / Toxicity Value) derived from CTV (left panels) or HTS assays (right panels), which is compared

to the MOE or HQ derived using regulatory toxicity values on the y-axis. Comparisons are made for regulatory NOAELs (panels A and B), BMDLs (panels C and D), or RfDs (panels E and F). In all cases, the predictions from CTV are based on cross-validation (panels A, C, and E). Each panel also includes lines indicating equality and 10-fold greater or less than equality (grey solid and dotted lines), nominal risk characterization thresholds (MOE = 100; HQ = 1), the number of compounds *n*, and the adjusted R^2 based on a linear model of log-transformed toxicity values. RfD = Reference Dose; NOAEL = No observed adverse effect level; BMDL = Benchmark dose lower confidence limit; OED05 = High throughput screening-based oral equivalent dose lower 5% confidence limit.

References

Crump KS, Chen C, Louis TA. 2010. The future use of in vitro data in risk assessment to set human exposure standards: Challenging problems and familiar solutions. Environ Health Perspect 118:1350-1354.

Sipes NS, Wambaugh JF, Pearce R, Auerbach SS, Wetmore BA, Hsieh JH, et al. 2017. An intuitive approach for predicting potential human health risk with the tox21 10k library. Environmental science & technology.

Wambaugh JF, Setzer RW, Reif DM, Gangwal S, Mitchell-Blackwood J, Arnot JA, et al. 2013. High-throughput models for exposure-based chemical prioritization in the expocast project. Environmental science & technology 47:8479-8488.

Wetmore BA. 2015. Quantitative in vitro-to-in vivo extrapolation in a high-throughput environment. Toxicology 332:94-101.