Published in final edited form as:
Arch Environ Contam Toxicol. 2022 November ; 83(4): 326–338. doi:10.1007/s00244-022-00944-5.

Relationships between aquatic toxicity, chemical hydrophobicity, and mode of action: log Kow revisited

Faith N. Lambert[1,3], **Deborah N. Vivian**[1], **Sandy Raimondo**[1], **Caroline T. Tebes-Stevens**[2], **Mace G. Barron**[1]

[1]Office of Research and Development, U.S. EPA, U.S. EPA, 1 Sabine Island Drive, Gulf Breeze, FL, USA

[2]Office of Research and Development, U.S. EPA, U.S. EPA, Athens, GA, USA

[3]Present Address: Syngenta, Research Triangle Park, NC 27709, USA

Abstract

Relationships between toxicity and chemical hydrophobicity have been known for nearly 100 years in mammals and fish, typically using the log of the octanol:water partition coefficient (Kow). The current study reassessed the influence of mode of action (MOA) on acute aquatic toxicity-log Kow relationships using a comprehensive database of 617 organic chemicals with curated and standardized acute toxicity data that did not exceed solubility limits, their consensus log Kow values, and weight of evidence-based MOA classifications (including 6 broad and 26 specific MOAs). A total of 166 significant ($p < 0.05$) log Kow-toxicity models were developed across six taxa groups that included QSARs for 5 of the broad and 13 of the specific MOAs. In this study, we demonstrate that QSARs based on MOAs significantly increase LC50 prediction accuracy for non-narcotics. Prediction accuracy increases when QSARs are built based on highly specific MOAs, rather than broad MOA classifications. Additionally, we demonstrate that building QSAR models with chemicals in specific MOA groupings, rather than broader MOA groups leads to significantly better estimates. We also evaluated the differences between models developed from mass-based (μg/L) and mole-based (μmol/L) toxicity data and demonstrate that both are suitable for QSAR development with no clear trend in greater model accuracy. Overall, the results reveal that, despite high variance in all taxa and MOA groups, specific MOA-based models can improve the accuracy of aquatic toxicity predictions over more general groupings.

INTRODUCTION

Relationships between chemical potency and hydrophobicity have been known for over 100 years in mammals and for decades in fish (Ferguson, 1939; Hansch & Dunn, 1972; Könemann, 1981; Levy & Gucinski, 1964; Lipnick, 1989; Meyer, 1899; Veith, Call, & Brooke, 1983). The log of the octanal/water partition coefficient (Kow) describes the ratio of a chemical's concentrations in the octanol and aqueous phases and is the most frequently used measure of chemical hydrophobicity. The log Kow is frequently referenced for toxicological profiling of chemicals because it is a readily obtainable physical-chemical

Corresponding author: Mace G. Barron, Barron.Mace@epa.gov.

property that highly correlates with acute toxicity and bioconcentration in tissues (McCarty, 1986). Accordingly, log Kow has become an established predictive descriptor in numerous quantitative structure activity relationship (QSAR) models of aquatic toxicity (Bradbury, Russom, Ankley, Schultz, & Walker, 2003).

Most toxicity-log Kow models have been developed for compounds of moderate polarity that exhibit baseline toxicity (M. G. Barron, Anderson, Lipton, & Dixon, 1997; McCarty, 1986; Schultz, Hewitt, Netzeva, & Cronin, 2007; Veith et al., 1983). Baseline toxicity, also known as narcosis, is a reversible, non-specific mechanism of action (MOA) that describes the minimal toxicity caused by a given chemical partitioning into biological membranes (M. G. Barron et al., 1997; McCarty, 1986). Previous aquatic toxicity-log Kow models for narcosis illustrate strong, positive correlations between increasing potency and hydrophobicity ($R^2 > 0.8$). However, many chemicals exhibit excess toxicity beyond the baseline mechanism through interactions with different molecular targets by specific MOAs. Although many contaminants characterized by mechanisms of excess toxicity, baseline toxicity models that do not account for differences in chemical MOA are still the most prominently characterized and used (e.g., (McCarty, 1986; Schultz et al., 2007; Veith et al., 1983)).

The field of aquatic toxicology has long recognized the importance of MOA in discriminating chemical potency. As such, some MOA-specific QSAR models have been developed for a variety of chemicals (e.g., (M. G. Barron et al., 1997; Bradbury et al., 2003; Escher & Hermens, 2002; Nendza & Muller, 2007; Netzeva, Pavan, & Worth, 2008; Russom, Bradbury, Broderius, Hammermeister, & Drummond, 1997; Zhang et al., 2013)). However, in the case of non-narcotics, toxicity is less well correlated with log Kow, necessitating the use of more computationally complex descriptors (Kar & Roy, 2010; Martin, Young, Lilavois, & Barron, 2015; Netzeva et al., 2008). Although more complex models potentially estimate toxicity accurately, they may be less accessible and interpretable as predictive tools. Understanding the role of MOA classifications with respect to log Kow-based QSAR development may inform best-practices for increasing accuracy of simple toxicity-log Kow models for compounds with excess toxicity.

Although standardized chemical MOA classifications have been established (Barron et al., 2015; Kienzler et al., 2019), no large-scale assessment has been reported that could better-inform QSAR model development. Evidence-based MOA classification frameworks are demonstrated to improve other toxicity estimation models such as interspecies correlation estimation (ICE) models (Raimondo et al., 2010) and may be equally useful for toxicity-log Kow QSARs. Barron et al. (2015) outlines a classification scheme that grouped chemicals with baseline and excess toxicity by broad MOAs as well as more specific mechanism-based MOA subcategories. Pairing this MOA scheme with QSAR model development could address several high-priority needs outlined by Cronin (2017), including exploring the applicability domain of non-polar narcosis and exploring potency by MOA for sensitive species.

Log Kow-based QSARs continue to be used to predict aquatic toxicity for both narocitic and more specific MOAs. Molar units have been recommended for aquatic toxicity QSAR

development, whereas pharmacology and medicinal chemistry have typically used mass-based concentration units (Dearden et al., 2009; Cherkasov et al., 2014). Recommendations for the use of molar units in QSARs have been based on either theoretical arguments of chemical effects or empirical assessments of goodness of fit. A rigorous statistical comparison across a range of chemicals, MOAs, and species has not been reported. The current study comprehensively and quantitatively determined the influence of MOA on aquatic toxicity-log Kow relationships by using multiple combinations of taxa and MOA groupings using both molar and mass-based QSARs. The assessment included four commonly tested aquatic organisms (three fish, one invertebrate) and combined taxa groups, six broad MOA groups (narcosis, non-narcotic MOAs) and twenty-six MOA subgroups that encompassed specific toxicity mechanisms. All analyses were conducted using the largest dataset of highly curated and standardized acute aquatic toxicity data to date, paired with consensus log Kow values and systematic MOA classifications (Barron et al., 2015). Notably, part of the data curation process involved verification and removal of toxicity values exceeding water solubility limits, which is not traditionally considered in aquatic toxicity model development but could affect model accuracy. Our results provide insight regarding best-practices for grouping chemical toxicity data when developing simple toxicity-log Kow QSAR models.

MATERIALS AND METHODS

Database development

The database was comprised of chemical name, chemical abstract registry number (CASRN), MOA classification, and values for log Kow, water solubility, and acute aquatic toxicity (median lethal concentration; fish LC50; invertebrate LC50 or immobilization EC50) based on standard international test guidance, or equivalent, for determining lethality (Barron et al., 2015). The final curated dataset of MOA classifications, and log Kow and toxicity values for 617 compounds is searchable by CASRN (Table SI–1). The toxicity data and MOA classification used in this study were obtained from MOAtox, which is a curated public domain database of acute toxicity values and MOA assignments (Barron et al., 2015). Acute toxicity data were comprised of both measured and estimated toxicity values that were highly standardized for experimental conditions. Estimated values were calculated using interspecies correlation estimation (ICE) models, which are empirical regressions of the toxicity values from a surrogate and predicted species (Raimando et al., 2010). MOA assignments were determined from the existing MOAtox classifications, which are based on a weight of evidence approach (Barron et al., 2015), and included 6 broad and 26 specific classes. All chemicals were categorized by a single broad MOA. Most chemicals were also categorized into a specific MOA category (29 chemicals were categorized as “other” for specific MOA and were not included in model development for groups of specific modes of toxic action). Table 1 lists the MOAs and outlines the subdivision of each broad class into more specific classes, including electron transport inhibition (ETI). Toxicity data for each chemical in MOAtox were binned into one of six taxa groups: *Daphnia* species (primarily *Daphnia magna*), bluegill (BG), fathead minnow (FHM), rainbow trout (RT), combined fish, and combined aquatic invertebrates. Organometallic and inorganic chemicals were excluded from the QSAR modeling set. Dissociating compounds were also excluded, except for a

few sodium (e.g., cyanide) and hydrochloride (e.g, formetanate) salts that were considered appropriate for log Kow modeling. Only records for organic chemicals that had associated toxicity data, MOA assignments, and log Kow values were retained.

Log Kow and water solubility values were compiled for all chemicals in the database, and measured values were used when available. Measured log Kow and solubility values were obtained from EpiSuite (www.epa.gov/tsca-screening-tools/epi-suitetm-estimation-program-interface; USEPA 2012) and/or PHYSPROP databases, a part of the EpiSuite software (SRC, North Syracuse, NY USA). Estimated Log Kow values were calculated for all chemicals using nine available predictive modeling tools for Kow: EPISuite KOWWIN (USEPA, 2012), SPARC (Hilal et al., 2004), ACD/Labs (https://www.acdlabs.com/products/percepta/predictors/logP/; ACD, Toronto, Ontario, Canada), OPERA (Mansouri et al., 2018), ChemAxon's calculator plugins (www.chemaxon.com/products/calculator-plugins/; ChemAxon, Cambridge, MA USA; using three methods: KLOP, PHYS and VG), and the OCHEM website (ochem.eu/home/show.do; using two methods: ALOGPS 2.1 and ALOGPS 3.0) (Sushko et al., 2011; Tetko & Tanchuk, 2002). The modeling tool with the lowest root mean square error (RMSE) between predicted and measured log Kow values was the ALOGPS 3.0 model (Tebes-Stevens et al., 2018). Therefore, ALGOPS3.0 log Kow estimates were used in the database when measured values were not available. Similarly, water solubility values were computed for all chemicals using EpiSuite WSKOW, EPISuite WATERNT, SPARC, ChemAxon's Calculator Plugin for solubility, ACD labs, OPERA, T.E.S.T. application (using four methods: group contribution, hierarchical, FDA, and nearest neighbor) (Martin et al., 2008), and the OCHEM website (using two methods: ALOGPS 2.1 and ALOGPS 3.0). The OPERA model produced the lowest RMSE between predicted and measured water solubility values. Therefore, the OPERA model solubility estimates were obtained from the EPA Chemistry dashboard (comptox.epa.gov/dashboard) and used in the database when measured values were not available. However, approximately 2% of chemicals in the database did not have available OPERA predictions. In these instances, the median predicted value from all other tools was used.

Quality assurance and curation

Quality control and quality assurance followed a two-step process in which (1) outliers were identified and evaluated and (2) toxicity data were compared to water solubility limits. To identify evaluate outliers, simple least squares regression models of the toxicity-log Kow relationships were developed for each combination of MOA and taxa group, and used to calculate 95% confidence intervals. Toxicity values that fell three orders of magnitude outside of the confidence limits were flagged as potential outliers. The original data source for each potential outlier was inspected and values that were not verifiable in the source document (e.g. by toxicity units, species identification, or test condition) were removed from the database. Toxicity values were then compared to the water solubility limit of each chemical, and those exceeding this threshold were removed from the dataset.

Model development

Log Kow-toxicity QSAR models ($n \geq 3$) were developed for different combinations of taxa (4 fish groups, 2 invertebrate groups) and MOA (6 broad groups and 26 specific groups).

QSAR models were developed using simple log-linear regression using the following equation:

$$log\ \ Toxicity = \beta \times log\ \ Kow + \alpha$$

where β and α represent the slope and intercept, respectively. Separate QSAR models were developed for mass-based (μg/L) and mole-based (μmol/L) toxicity data. Only significant models ($p < 0.05$) were used for further analyses. Model parameters, including R^2, slope, intercept, and mean square error (MSE), were also computed (Table SI–2).

Data analyses

Model validation.—The prediction accuracy of models containing four or more points was assessed using leave-one-out cross-validation (LOOCV). In LOOCV, each data point, which represents a paired log acute toxicity value (LC50) and log Kow value, was systematically removed from the original model. Then, a new model (sub-model) was built with the remaining data. Each sub-model was used to estimate the removed LC50 using its associated log Kow value. LOOCV was performed on all models with $n \geq 4$. The fold difference of each predicted and measured LC50 (using non-transformed data) was calculated as estimated/actual or actual/estimated (whichever was larger) and used as the validation metric (Raimondo et al., 2010). Additionally, the prediction accuracy was calculated as the estimated/actual LC50 (using non-transformed data) to preserve directionality for illustrative purposes and to visually confirm non-systematic variance of models.

In addition to the LOOCV for broad and specific MOA models, we evaluated the use of baseline toxicity models (i.e., those developed using data for chemicals with a broad MOA of narcosis) for estimating toxicity of chemicals with excess toxicity (i.e., non-narcotics). The log Kow of each non-narcotic chemical was used to estimate toxicity from each species-specific baseline toxicity model. For each estimate, the fold difference from the measured toxicity value and prediction accuracy were calculated as described above.

Comparison of μg/L and μmol/L models.—An analysis was conducted to determine whether model quality differed significantly between QSARs built using μg/L or μmol/L toxicity data. To achieve this, a pairwise comparison of model MSEs between QSARs built using μg/L and μmol/L data was conducted. The MSE is the sum of squared differences between the predicted and actual values of a model, thereby representing a model's fit to the data. MSE values are strictly positive, with smaller values indicating a better fit (a value of zero represents a perfect fit). The MSE of μg/L and μmol/L models were paired by taxa and MOA and compared using a Wilcoxon signed rank test. Separate analyses were conducted for broad and specific MOA models.

Comparison of baseline toxicity, broad MOA, and specific MOA models.—Three analyses were conducted to assess (1) the prediction accuracy of simple baseline toxicity models for chemicals with excess toxicity, (2) differences in prediction accuracy of broad or specific MOA models for the same compounds, and (3) whether any individual broad

or specific MOA models performed better than others. Both analyses were done separately for models built with mass-based (μg/L) and mole-based (μmol/L) toxicity data. First, to assess differences between broad and specific MOA models, we conducted a pairwise comparison of toxicity predictions. The fold-difference of predicted toxicity values for broad and specific MOA models were paired by taxa and CAS number and evaluated using a Wilcoxon signed rank test. Second, we conducted two-way ANOVAs to determine whether individual MOA classes performed significantly better or worse than others (broad and specific MOAs were evaluated separately). Each two-way ANOVA compared the mean fold-difference of predicted values with MOA as the treatment factor and taxa as the blocking factor.

Potency differences across broad and specific MOAs.—We assessed how grouping by broad/specific MOAs could isolate chemical groups with distinct potency profiles. Two analyses were conducted to determine if potency (1) differed between broad MOAs and (2) differed between specific MOAs within a broad MOA group. Analyses were conducted on log-transformed toxicity values from models built with mass-based data (μg/L). Significant differences were assessed with one-way ANOVAs followed by pairwise t-tests with Bonferroni correction. Analyses were conducted separately for one representative fish (rainbow trout) and invertebrate (*D. magna*).

RESULTS

Database

The final curated database used in developing QSARs was comprised of aquatic toxicity data for 617 organic chemicals, accompanied by consensus log Kow values and MOA assignments (Table SI–1). Log Kow values spanned the range of −5.4 to 8.33 across all chemicals. Across all taxa groups, non-polar (specific MOA) narcotics (broad MOA) was the largest chemical group within the database (n=152 to 262 non-polar narcotic chemicals per taxon) (Tables 1 and 2).

Model development

A total of 166 significant ($p < 0.05$) log Kow-toxicity models were developed, which represented all 6 taxa groups, 5 of 6 broad MOAs, and 13 of 26 specific MOAs in the complete data set (Table 2, Table SI–2). Despite a quantitative and systematic outlier removal process, large scatter in the data were observed for all MOA and taxa groups (Figs. SI–1, SI–2, SI–3, SI–4, SI–5, SI–6, SI–7, SI–8, SI–9, SI–10, SI–11, SI–12). The ranges of significant model parameters are outlined in Table 2. Prediction accuracy showed non-directional variance across all MOAs and taxa groups (Fig. 1). On average, log Kow explained 59% and 58% of the variation in broad and specific MOA models, respectively.

There were slightly more significant mole-based models (87 total; Tables 1, SI–3) than mass-based models (79 total; Tables 1, 2). This prompted us to evaluate differences in MSE between the two groups to determine whether mole-based models had a significantly better-fit to the data. Although we expected MSEs of the more abundant mole-based models to indicate a better fit, we found the opposite. The average MSE for QSARs developed using

mass-based toxicity data (μg/L) was significantly lower than that of mole-based (μmol/L) models, for both broad (V statistic 25, p = 1.68e-06) and specific models (V statistic 72, p = 1.31e-09) (Fig. 2). When graphed, mole- and mass-based models had visibly similar scatter (Fig. 3).

Comparison of prediction accuracy using baseline, broad MOA, and specific MOA models

QSAR prediction accuracy significantly increased when models were built using chemicals with the same MOA. This suggests that grouping chemicals by MOA accounts for some of the variability in toxicity-log Kow QSARs, and that log Kow can be a more relevant predictor of toxicity for chemicals within a single MOA class. The percentage of cross-validated data points that were predicted within 5-fold of the measured value ranged from 50.5 to 70.4% across broad MOA groupings, and from 49.0–100% for specific MOA groupings (Table 2). Overall, the fold-difference of predicted from measured toxicity values was significantly smaller for specific MOAs than broad MOAs (V Statistic = 1993742, p-value = 6.55e-10), indicating that specific MOA QSARs had better prediction accuracy (Fig. 4). Analysis of the average fold-difference of predictions across different MOAs indicated that no single broad or specific MOA QSAR performed better than any other (Fig. 5). The lowest prediction accuracies occurred when baseline QSAR models were used to predict toxicity of chemicals with excess toxicity. Only 0.0 – 52.8% of predictions fell within 5-fold of the measured values (Table 2), and toxicity was generally underestimated (Fig. 5). Mass-based results are presented in the main text and results for mole-based models can be found in the SI (Table SI–3, Figs SI–13, SI–14).

Comparison of potency across MOAs

We analyzed potency differences between MOAs in a representative fish (rainbow trout) and invertebrate (*D. magna*). In general, broad MOA groups had distinct potencies (Fig. 6) for both taxa. Notably, the overall potency of narcotics (baseline toxicity) was significantly less than all other MOAs (exhibiting excess toxicity) (Fig. 6). Subdividing chemicals further into specific MOA groupings resulted in additional groups with distinct potencies (Fig. 7).

DISCUSSION

Historically, aquatic toxicity QSARs have been developed with log Kow as the only descriptor and have been highly predictive of chemicals with baseline (minimal) toxicity that is determined by chemical partitioning and a narcosis MOA (Könemann, 1981; Levy & Gucinski, 1964; Lipnick, 1989; Veith et al., 1983). More recently, log Kow QSARs have been applied to compounds with diverse and often complex chemical structures and with specific mechanisms of toxicity including pesticide neurotoxicants, reactive nucleophiles, and electron transport inhibitors (e.g., Mayo-Bean et al., 2012). The current study is the first to systematically evaluate the prediction accuracy of log Kow QSARs across a diverse range of MOA and aquatic species groupings. We utilized the largest dataset of highly curated and standardized acute toxicity data, consensus log Kow values, and systematic MOA classifications ever applied in aquatic toxicity QSAR modeling to address the extent to which MOA categorization could improve QSAR models (Barron et al., 2015). The overall

results of these analyses showed that baseline toxicity models had generally lower prediction accuracy for compounds with excess toxicity than MOA-specific QSARs.

It is commonly accepted that aquatic toxicity and log Kow can be poorly correlated for specifically acting chemicals, and therefore mechanism-based models have been developed using more complex descriptors and model formulations (e.g., Barron et al., 1997; Bradbury et al., 2003; Escher & Hermens, 2002; Lee & Barron, 2015; Netzeva et al., 2008; Russom et al., 1997; Zhang et al., 2013). Early development of QSARs utilized steric (size, shape) and electronic factors in addition to hydrophobicity as a predictors of chemical toxicity (e.g., Hansch et al., 1991), and more recent research has advanced the use of more complex formulations (e.g., Böhme et al., 2016). However, international QSAR tools such as ECOSAR continue to use log kow to predict toxicity of both neutral organics and more complex structures (e.g., Mayo-Bean et al. 2012). In the current study, QSARs developed only using log Kow had generally better prediction accuracy with increasing specificity of the MOA category. This observation may be attributed to the distinct mechanisms and potencies within different chemical groups. Chemicals in broad MOA categories had distinct toxicity ranges within taxa. When subdivided into specific MOA groups, more chemical groups with distinct potencies were apparent (e.g., for acetylcholinesterase inhibition, neurotoxicity). These differences are likely related to discrete mechanisms of toxicity and ranges of organism susceptibility. Parsing out such differences through increasingly narrow MOA classifications appears to decrease model error. However, despite the increased accuracy, log Kow explained a similar amount of variance. On average, log Kow accounted for only 59% (R^2 = 0.171–0.867) and 54% (R^2 = 0.124–0.994) of broad and specific MOA models, respectively. Other studies have reported smaller variance and higher R^2 for log Kow based QSARs of specific acting chemicals for generally narrow ranges of chemicals using either MOA or chemical structure-based compound groupings (Cronin, 2017; Mayo-Bean et al., 2012; Netzeva et al., 2008). The results demonstrate variable accuracy in log Kow QSARs and suggest how existing MOA schemes can be used to inform model development and improve prediction accuracy.

Narcotics continue to be the largest class of chemicals with aquatic toxicity data (Barron et al., 1997; Kienzler et al., 2017; Kienzler et al., 2019; Veith et al., 1983; Zhang et al., 2013), and they comprised 60% of the 617 chemicals in the current database across all taxa groupings. Broad narcosis models were statistically significant for each of the six taxa groups, but an average of only 50% of the variance in toxicity was explained by log Kow (R^2 = 0.311–0.583). An assessment of the performance of six international QSAR tools using a dataset of 130 chemicals dominated by narcotics also showed limited prediction accuracy, with only 44 to 55% of model estimates within a factor of 5 of measured values (Moore et al., 2003). The structural complexity of the diversity of chemicals in the datasets likely contributed to lower performance of QSAR models in the current and past studies (e.g., Moore et al., 2003). However, as noted for other chemical groups, the use of more narrow MOA categories for narcotics (polar, nonpolar, ester) had a positive effect on prediction accuracy and, in some cases, accounted for more of the model variance. In specific narcotic models, log Kow accounted for an average of 67% (R^2 = 0.171–0.867), 49% (R^2 = 0.353–0.565), and 59% (R^2 = 0.234–0.775) for polar, nonpolar, and ester subgroups, respectively. Therefore, narcotic chemical QSARs can also benefit from narrower constraints.

The current study is the first to comprehensively compare the prediction accuracy of mass-based and mole-based across a wide variety of chemicals exhibiting baseline and excess toxicity. Narcosis-based QSARs have been traditionally developed using molar units of toxicity, and molar units have been recommended in model development by Cherkasov et al. (2014), Dearden et al. (2009), and others. In contrast, pharmacology and medicinal chemistry studies have traditionally used mass-based units, based on the theory of mass-action that predicts the interaction of a drug and receptor is determined by the amount of chemical present at the site of action. Aquatic toxicologists primarily utilize mole-based units for QSAR modeling (Barron et al., 1997; McCarty, 1986; Schultz et al., 2007; Veith et al., 1983) based on theories that the number of molecules rather than mass at the target site drives toxicity. In the current study, although more of the developed mole-based (μmol/L) models were significant, mass-based (μg/L) models had better fit to the data as indicated by a lower MSE, and both models exhibited similar amounts of data scatter. These results suggest use of either mole- or mass-based models may be suitable for log Kow-based QSAR development, and potentially less dependent on the MOA classification. One possible reason for the similarity in molar and mass based QSARs may be the generally narrow range in molecular size across MOA and taxa groupings (e.g., 200–400 daltons). High correlations between mass and mole-based toxicity values over this narrow range would also suggest similarity in models.

MOA classification in aquatic toxicology can be uncertain because chemicals may act through multiple mechanisms depending on chemical substructure, exposure regime and taxon-specific disposition (Barron et al., 2015; Cronin, 2017; Escher & Hermens, 2002). Additionally, there is often a lack of concordance in MOA assignments across international classification schemes, and incorrect MOA assignments exceeding 50% of test compounds can occur with even simplified schemes (e.g., Cherkasov et al., 2014); Ellison, Madden, Cronin, and Enoch, 2015; Nendza and Muller, 2007). MOA can also shift within a class of structurally similar compounds through changes in substituents and moiety or alterations in chemical speciation from changes in test conditions (Escher & Hermens, 2002; von der Ohe et al., 2005). A classic example is the pH dependent toxicity of a homologous series of chlorophenols, where MOA changes from narcosis to electron transport inhibition (ETI) with increasing halogen substitution (M. G. Barron et al., 1997; Könemann & Musch, 1981). Misclassification of MOA can lead to large errors in QSAR-based toxicity estimation (Ellison et al., 2015; Nendza & Muller, 2007). MOA assignments in the current study were considered to have high certainty because they were based on only acute toxicity mechanisms, and used existing weight of evidence assignments in the MOAtox database (Barron et al., 2015), rather than model predictions. Any ambiguous MOA classification was excluded in the original MOAtox database (Barron et al., 2015), and verified in the dataset curation in the current study.

There are multiple challenges to developing MOA-based QSARs, including issues of data quality and availability, appropriate use of descriptors, and model statistics and validation common to all QSARs (Cronin & Schultz, 2003; Dearden et al., 2009). The availability of high-quality toxicity data for a diversity of species and chemicals has been a continuing limitation of QSAR development in aquatic toxicology. The current study utilized the extensively curated MOAtox dataset that included values estimated with ICE models to

provide for a greater number and diversity of chemicals and taxa in the dataset (Barron et al., 2015). As noted by Raimondo et al. (2010) and Willming, Lilavois, Barron, and Raimondo (2016), interspecies estimates are computed from validated toxicity models that are based on intrinsic species relationships and that are not predicted based on chemical structure or molecular properties. High prediction accuracy of ICE models has been demonstrated to be within the range of uncertainty inherent in toxicity data collected from multiple test laboratories. Supplementing the toxicity dataset with a limited number of ICE estimated values in the current study contributed minimal additional uncertainty while expanding the domain of applicability to additional chemical and taxa.

In conclusion, this study demonstrated that toxicity-log Kow QSARs based on specific MOA categories can improve prediction accuracy for a range of taxa and compounds. Results support the long-standing premise that log Kow is a fundamental physical property correlated with organic chemical toxicity across a variety of MOAs (Cherkasov et al., 2014; Cronin, 2017). We suggest that currently established MOA classification schemes can serve as tools to inform QSAR development and improve prediction accuracy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

REFERENCES

Barron MG, Anderson MJ, Lipton J, & Dixon DG. (1997). Evaluation of critical body residue QSARs for predicting organic chemical toxicity to aquatic organisms. SAR QSAR Environ Res, 6(1–2), 47–62. doi:10.1080/10629369708031724 [PubMed: 9241865]

Barron MG, Lilavois CR, & Martin TM. (2015). MOAtox: A comprehensive mode of action and acute aquatic toxicity database for predictive model development. Aquatic Toxicology, 161, 102–107. [PubMed: 25700118]

Böhme A, Laqua A, Schüürmann G. (2016). Chemoavailability of Organic Electrophiles: Impact of Hydrophobicity and Reactivity on Their Aquatic Excess Toxicity. Chem. Res. Toxicol 29, 952–962. [PubMed: 27096880]

Bradbury SP, Russom CL, Ankley GT, Schultz TW, & Walker JD. (2003). Overview of data and conceptual approaches for derivation of quantitative structure-activity relationships for ecotoxicological effects of organic chemicals. Environ Toxicol Chem, 22(8), 1789–1798. doi:10.1897/01-234 [PubMed: 12924578]

Cherkasov A, Muratov EN, Fourches D, Varnek A, Baskin II, Cronin M, Tropsha A. (2014). QSAR modeling: where have you been? Where are you going to? J Med Chem, 57(12), 4977–5010. doi:10.1021/jm4004285 [PubMed: 24351051]

Cronin MT. (2017). (Q)SARs to predict environmental toxicities: current status and future needs. Environ Sci Process Impacts, 19(3), 213–220. doi:10.1039/c6em00687f [PubMed: 28243641]

Cronin MTD, & Schultz TW. (2003). Pitfalls in QSAR. Journal of Molecular Structure: THEOCHEM, 622(1–2), 39–51. doi:10.1016/s0166-1280(02)00616-4

Dearden JC, Cronin MT, & Kaiser KL. (2009). How not to develop a quantitative structure-activity or structure-property relationship (QSAR/QSPR). SAR QSAR Environ Res, 20(3–4), 241–266. doi:10.1080/10629360902949567 [PubMed: 19544191]

Ellison CM, Madden JC, Cronin MT, & Enoch SJ. (2015). Investigation of the Verhaar scheme for predicting acute aquatic toxicity: improving predictions obtained from Toxtree ver. 2.6. Chemosphere, 139, 146–154. doi:10.1016/j.chemosphere.2015.06.009 [PubMed: 26092094]

Escher BI, & Hermens JL. (2002). Modes of action in ecotoxicology: their role in body burdens, species sensitivity, QSARs, and mixture effects. Environ Sci Technol, 36(20), 4201–4217. doi:10.1021/es015848h [PubMed: 12387389]

Ferguson J. (1939). The use of chemical potentials as indices of toxicity. Proc. Roy. Soc. B127, 387–404.

Hansch C, & Dunn WJ 3rd. (1972). Linear relationships between lipophilic character and biological activity of drugs. J Pharm Sci, 61(1), 1–19. doi:10.1002/jps.2600610102 [PubMed: 4550859]

Hansch C, Leo A, Taft RW. (1991). A survey of Hammett substituent constants and resonance and field parameters. Chemical reviews 91, 165–195.

Hilal SH, Karickhoff SW, & Carreira LA. (2004). Prediction of the Solubility, Activity Coefficient and Liquid/Liquid Partition Coefficient of Organic Compounds. QSAR & Combinatorial Science, 23(9), 709–720. doi:10.1002/qsar.200430866

Kar S, & Roy K. (2010). QSAR modeling of toxicity of diverse organic chemicals to Daphnia magna using 2D and 3D descriptors. J Hazard Mater, 177(1–3), 344–351. doi:10.1016/j.jhazmat.2009.12.038 [PubMed: 20045248]

Kienzler A, Barron MG, Belanger SE, Beasley A, & Embry MR. (2017). Mode of Action (MOA) Assignment Classifications for Ecotoxicology: An Evaluation of Approaches. Environ Sci Technol, 51(17), 10203–10211. doi:10.1021/acs.est.7b02337 [PubMed: 28759717]

Kienzler A, Connors KA, Bonnell M, Barron MG, Beasley A, Inglis CG, . . . Embry MR. (2019). Mode of Action Classifications in the EnviroTox Database: Development and Implementation of a Consensus MOA Classification. Environ Toxicol Chem, 38(10), 2294–2304. doi:10.1002/etc.4531 [PubMed: 31269286]

Könemann H. (1981). Quantitative structure-activity relationships in fish toxicity studies Part 1: Relationship for 50 industrial pollutants. Toxicology, 19(3), 209–221. doi:10.1016/0300-483x(81)90130-x [PubMed: 7233445]

Könemann H, & Musch A. (1981). Quantitative structure-activity relationships in fish toxicity studies Part 2: The influence of pH on the QSAR of chlorophenols. Toxicology, 19(3), 223–228. doi:10.1016/0300-483x(81)90131-1 [PubMed: 7233446]

Lee S, & Barron MG. (2015). Development of 3D-QSAR Model for Acetylcholinesterase Inhibitors Using a Combination of Fingerprint, Molecular Docking, and Structure-Based Pharmacophore Approaches. Toxicol Sci, 148(1), 60–70. doi:10.1093/toxsci/kfv160 [PubMed: 26202430]

Levy G, & Gucinski SP. (1964). Studies on biologic membrane permeation kinetics and acute toxicity of drugs by means of goldfish. Journal of Pharmacology and Experimental Therapeutics, 146(1), 80–86. [PubMed: 14221231]

Lipnick RL. (1989). Structure-Activity relationships in environmental toxicology and chemistry: Narcosis, electrophile and proelectrophile toxicity mechanisms: Application of SAR and QSAR. Env Tox Chem, 8, 1–12.

Mansouri K, Grulke CM, Judson RS, & Williams AJ. (2018). OPERA models for predicting physicochemical properties and environmental fate endpoints. J Cheminform, 10(1), 10. doi:10.1186/s13321-018-0263-1 [PubMed: 29520515]

Martin TM, Harten P, Venkatapathy R, Das S, & Young DM. (2008). A hierarchical clustering methodology for the estimation of toxicity. Toxicol Mech Methods, 18(2–3), 251–266. doi:10.1080/15376510701857353 [PubMed: 20020919]

Martin TM, Young DM, Lilavois CR, & Barron MG. (2015). Comparison of global and mode of action-based models for aquatic toxicity. SAR QSAR Environ Res, 26(3), 245–262. doi:10.1080/1062936X.2015.1018939 [PubMed: 25783870]

Mayo-Bean K, Moran K, Meylan B, & Ranslow P. (2012). Methodology document for the ecological structure-activity relationship model (ECOSAR) class program. Estimating toxicity of industrial chemicals to aquatic organisms using the ECOSAR (ecological structure activity relationship) class program. . In (pp. 46): MS-Windows Version 1.1.11. .

McCarty LS. (1986). The relationship between aquatic toxicity QSARs and bioconcentration for some organic chemicals. Environmental Toxicology and Chemistry, 5(12), 1071–1080. doi:10.1002/etc.5620051207

Meyer H. (1899). The theory of alcohol anesthesia. Arch Exp Path Pharm, 42, 109–118.

Moore DR, Breton RL, & MacDonald DB. (2003). A comparison of model performance for six quantitative structure-activity relationship packages that predict acute toxicity to fish. Environ Toxicol Chem, 22(8), 1799–1809. doi:10.1897/00-361 [PubMed: 12924579]

Nendza M, & Muller M. (2007). Discriminating toxicant classes by mode of action: 3. Substructure indicators. SAR QSAR Environ Res, 18(1–2), 155–168. doi:10.1080/10629360601054354 [PubMed: 17365966]

Netzeva TI, Pavan M, & Worth AP. (2008). Review of (Quantitative) Structure–Activity Relationships for Acute Aquatic Toxicity. QSAR & Combinatorial Science, 27(1), 77–90. doi:10.1002/qsar.200710099

Raimondo S, Jackson CR, & Barron MG. (2010). Influence of taxonomic relatedness and chemical mode of action in acute interspecies estimation models for aquatic species. Environ Sci Technol, 44(19), 7711–7716. doi:10.1021/es101630b [PubMed: 20795664]

Russom CL, Bradbury SP, Broderius SJ, Hammermeister DE, & Drummond RA. (1997). Predicting modes of toxic action from chemical structure: Acute toxicity in the fathead minnow (Pimephales promelas). Environmental Toxicology and Chemistry, 16(5), 948–967. doi:10.1002/etc.5620160514

Schultz TW, Hewitt M, Netzeva TI, & Cronin MTD. (2007). Assessing Applicability Domains of Toxicological QSARs: Definition, Confidence in Predicted Values, and the Role of Mechanisms of Action. QSAR & Combinatorial Science, 26(2), 238–254. doi:10.1002/qsar.200630020

Sushko I, Novotarskyi S, Korner R, Pandey AK, Rupp M, Teetz W, . . . Tetko IV. (2011). Online chemical modeling environment (OCHEM): web platform for data storage, model development and publishing of chemical information. J Comput Aided Mol Des, 25(6), 533–554. doi:10.1007/s10822-011-9440-2 [PubMed: 21660515]

Tebes-Stevens C, Patel JM, Koopmans M, Olmstead J, Hilal SH, Pope N, . . . Wolfe K. (2018). Demonstration of a consensus approach for the calculation of physicochemical properties required for environmental fate assessments. Chemosphere, 194, 94–106. doi:10.1016/j.chemosphere.2017.11.137 [PubMed: 29197820]

Tetko IV, & Tanchuk VY. (2002). Application of associative neural networks for prediction of lipophilicity in ALOGPS 2.1 program. J Chem Inf Comput Sci, 42(5), 1136–1145. doi:10.1021/ci025515j [PubMed: 12377001]

USEPA. (2012). Estimation Programs Interface Suite™ for Microsoft® Windows, v 4.11. . In. Washington, DC, USA: United States Environmental Protection Agency.

Veith GD, Call DJ, & Brooke LT. (1983). Structure–Toxicity Relationships for the Fathead Minnow, Pimephales promelas: Narcotic Industrial Chemicals. Canadian Journal of Fisheries and Aquatic Sciences, 40(6), 743–748. doi:10.1139/f83-096

von der Ohe PC, Kuhne R, Ebert RU, Altenburger R, Liess M, & Schuurmann G. (2005). Structural alerts--a new classification model to discriminate excess toxicity from narcotic effect levels of organic compounds in the acute daphnid assay. Chem Res Toxicol, 18(3), 536–555. doi:10.1021/tx0497954 [PubMed: 15777094]

Willming MM, Lilavois CR, Barron MG, & Raimondo S. (2016). Acute Toxicity Prediction to Threatened and Endangered Species Using Interspecies Correlation Estimation (ICE) Models. Environ Sci Technol, 50(19), 10700–10707. doi:10.1021/acs.est.6b03009 [PubMed: 27585402]

Zhang X, Qin W, He J, Wen Y, Su L, Sheng L, & Zhao Y. (2013). Discrimination of excess toxicity from narcotic effect: comparison of toxicity of class-based organic chemicals to Daphnia magna and Tetrahymena pyriformis. Chemosphere, 93(2), 397–407. doi:10.1016/j.chemosphere.2013.05.017 [PubMed: 23786811]

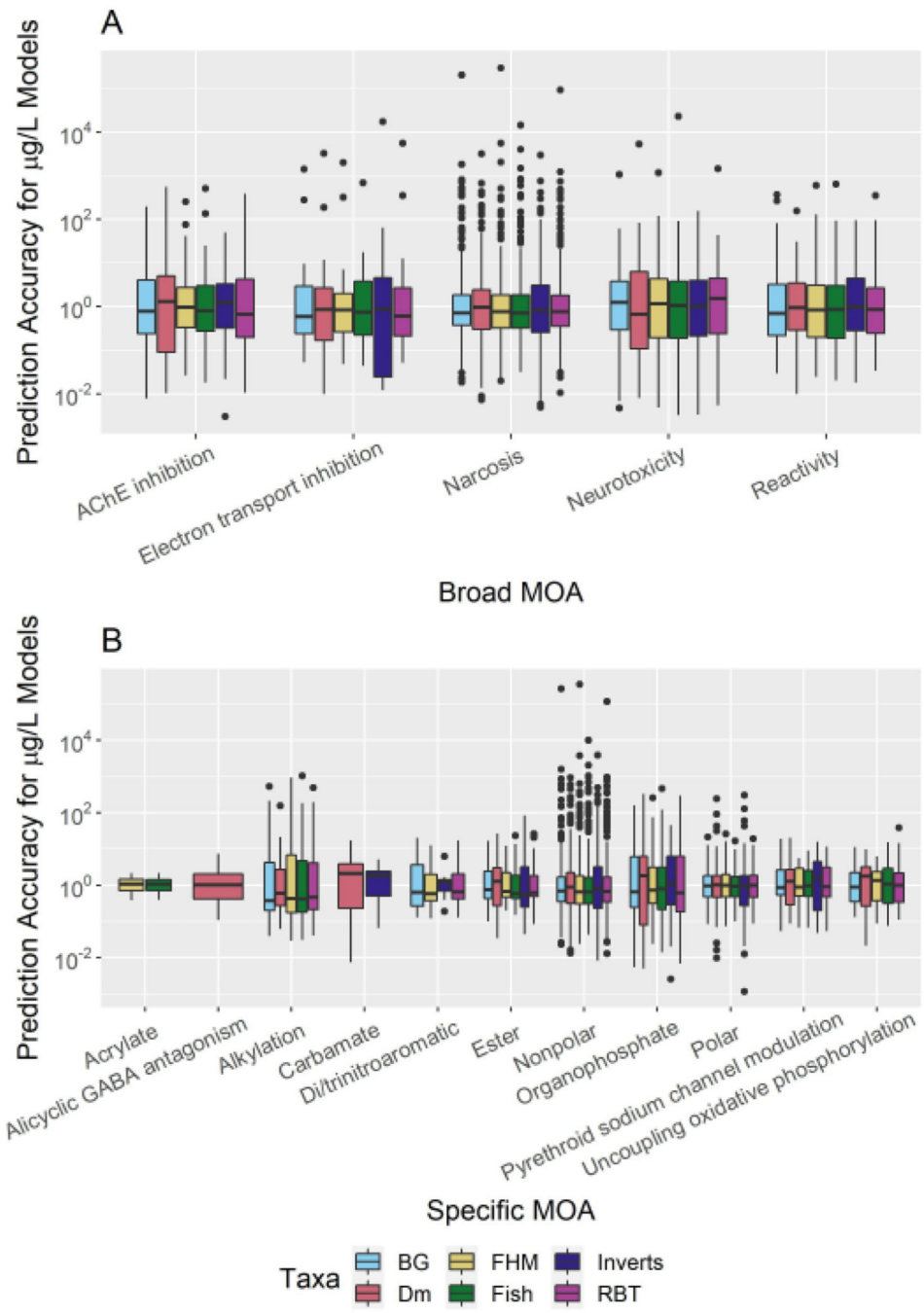

Figure 1.

Box plots of the prediction accuracy based on MOA and taxa for mass-based (µg/L) data. The top graph depicts data for broad MOA models and the bottom for specific MOA models. Predictions were non-systematic in all models. No broad or specific-based model performed better than any other in its category for any taxa.

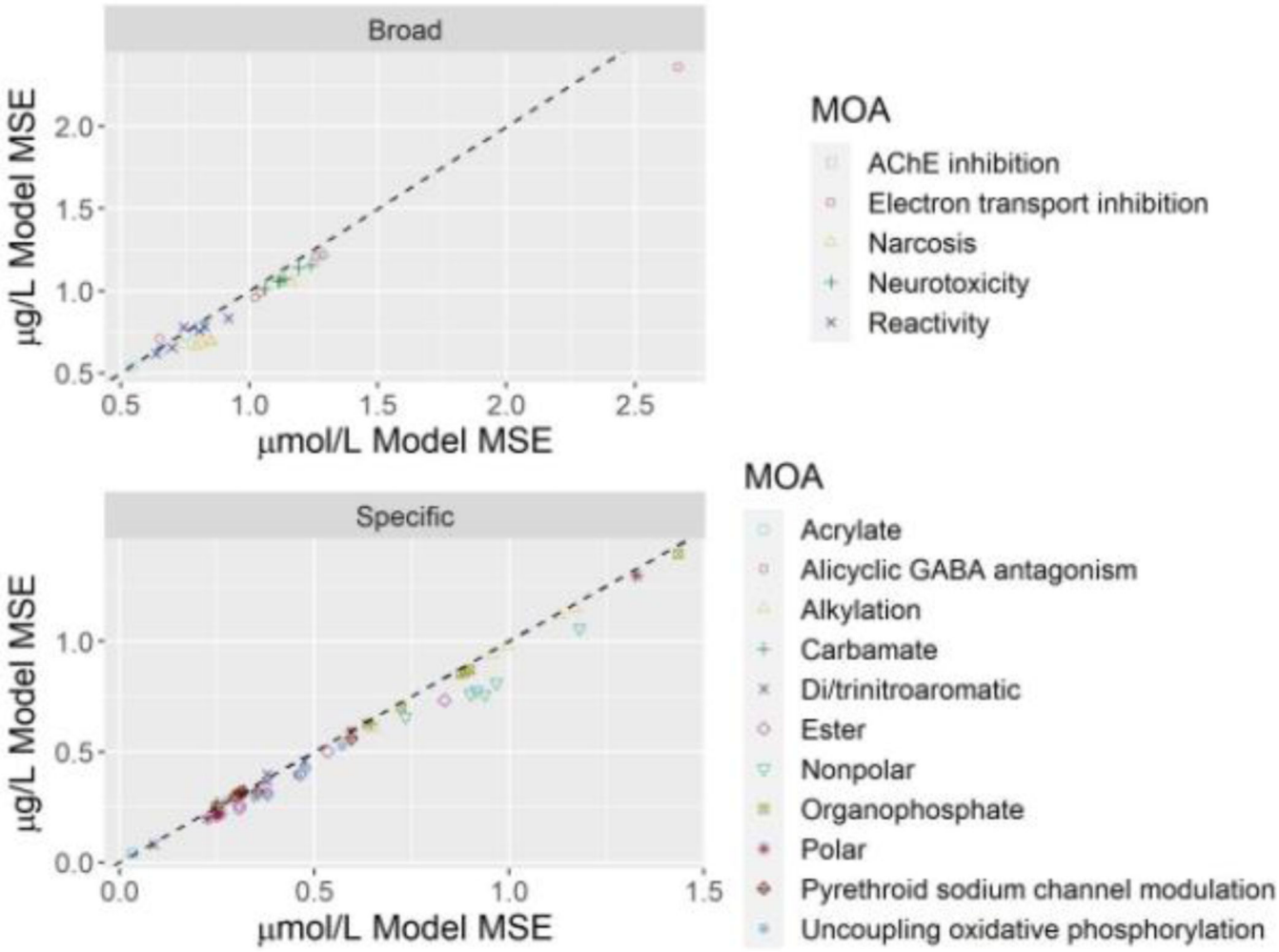

Figure 2.

Comparison of MSE between models made using mass-based (μg/L) or mole-based (μmol/L) toxicity data. Separate graphs were made for broad (top) and specific (bottom) MOA models. Data points are colored by the MOA class they represent. The dashed line is the 1:1 line.

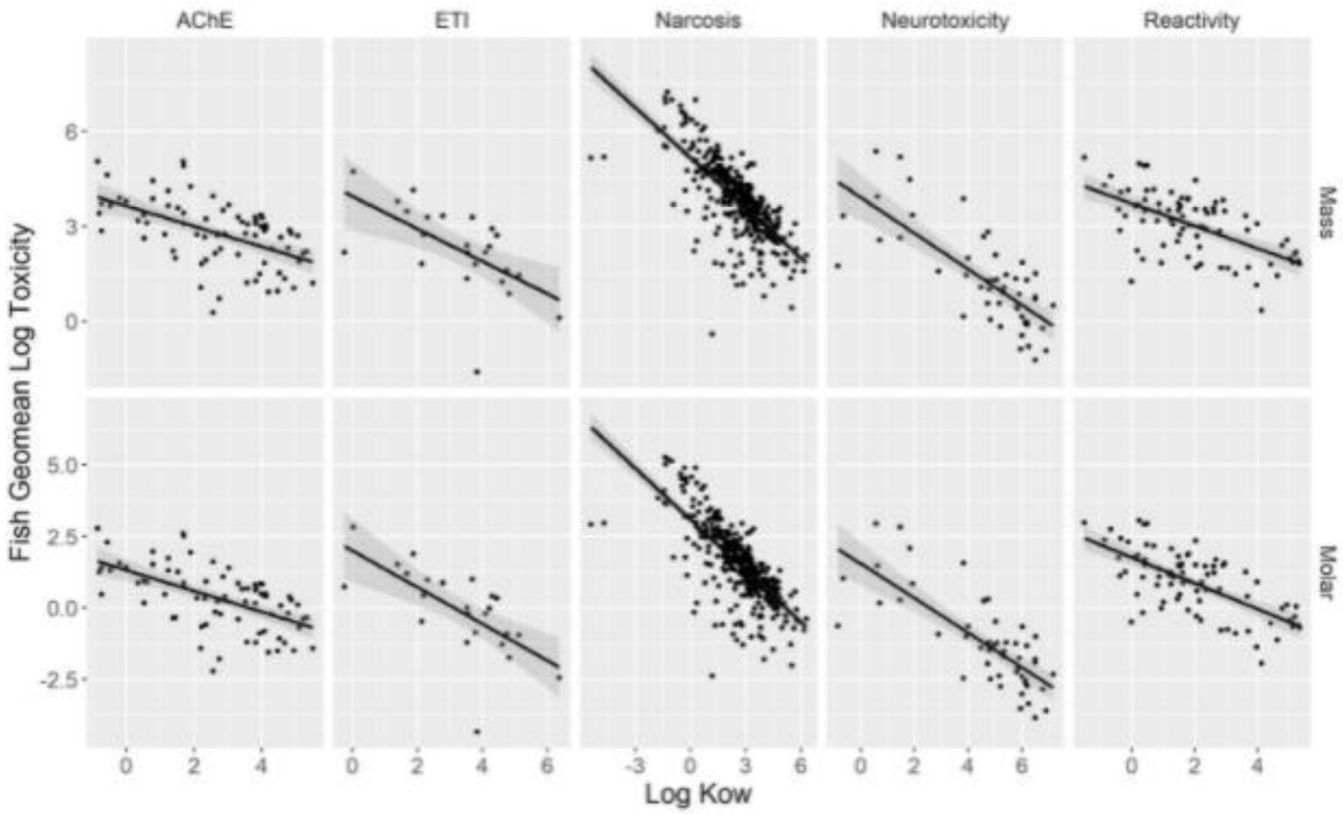

Figure 3.

Rainbow trout broad MOA models for mass-based (μg/L; top panels) and mole-based (μmol/L; bottom panels) toxicity data. Abbreviations: AChE = Acetylcholinesterase inhibition; ETI = Electron transport inhibition.

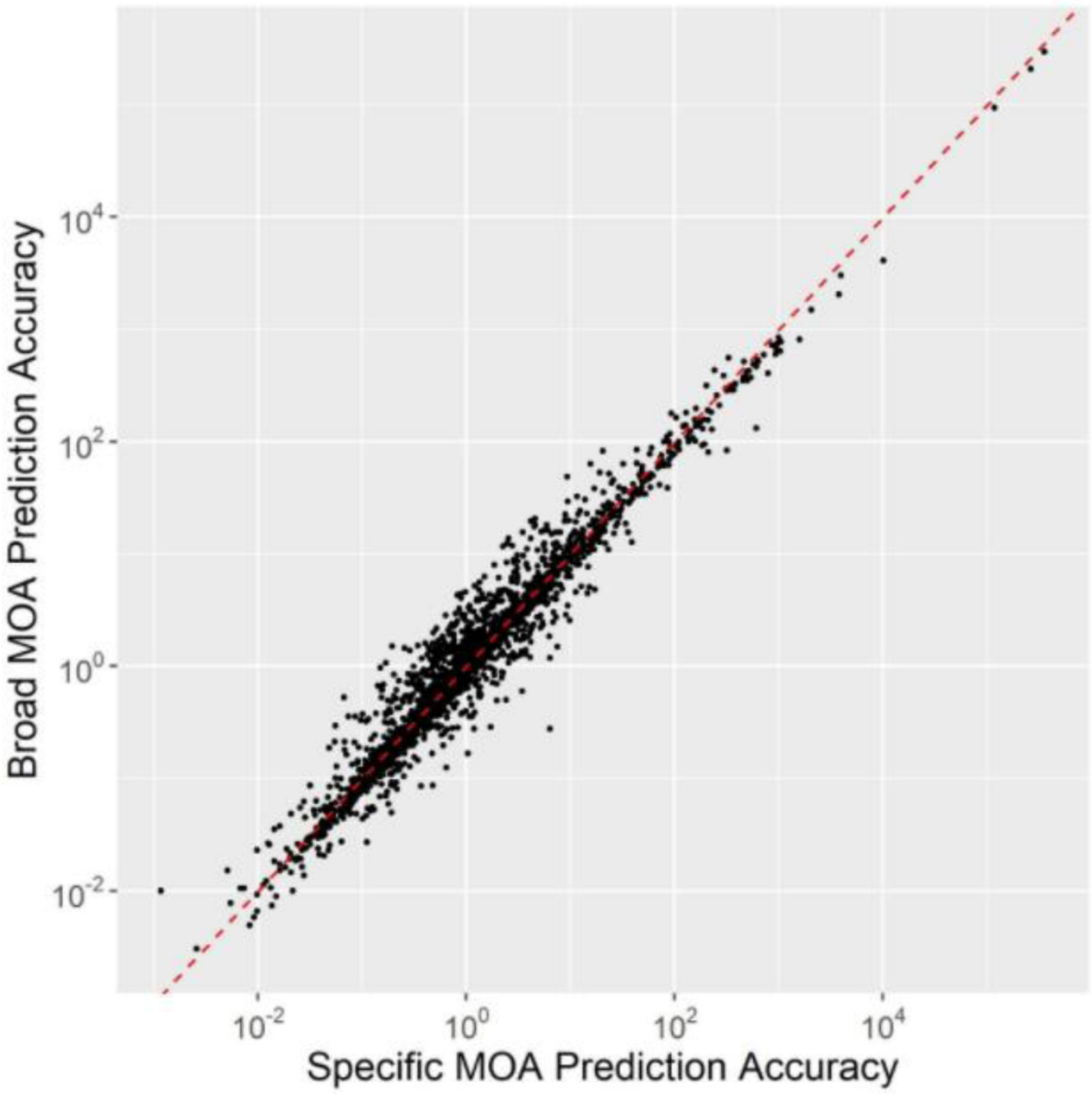

Figure 4.
Comparison of prediction accuracy of individual chemical toxicity between broad and specific MOA models. The dashed red line is the 1:1 line.

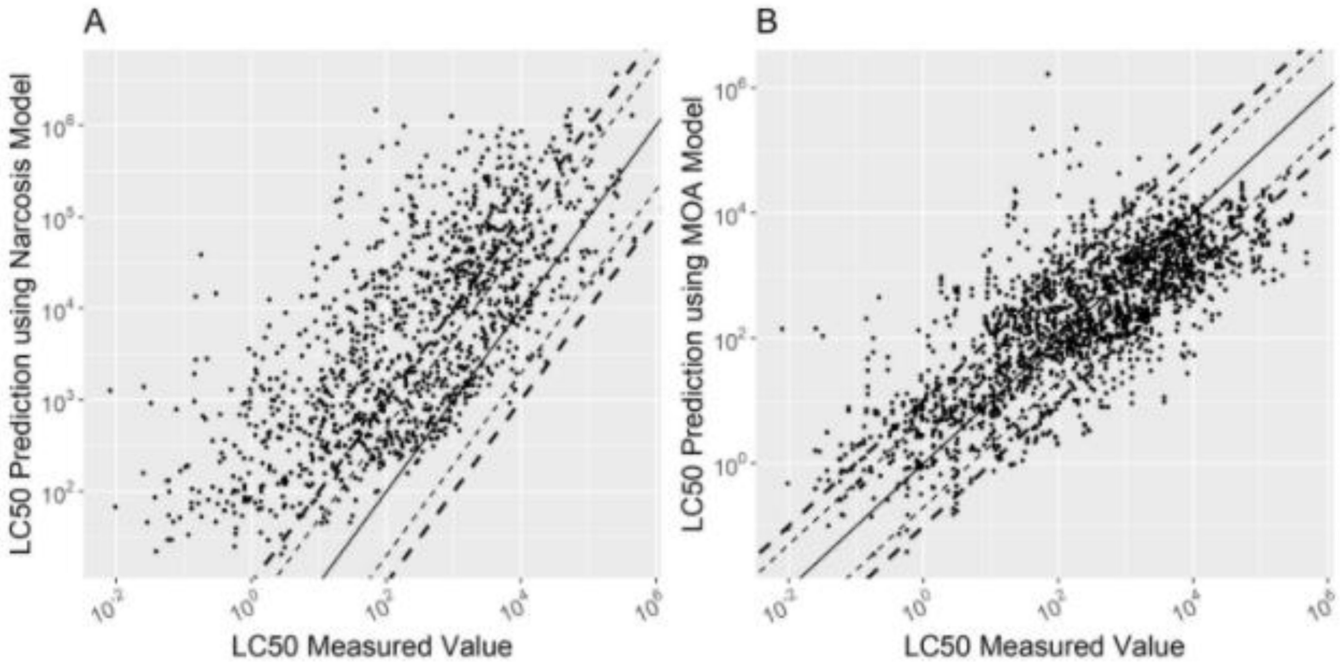

Figure 5.

LC50 predictions from log Kow values of chemicals exhibiting excess toxicity using (A) baseline toxicity models and (B) either broad or specific MOA models. The solid line is the 1:1 line, the thin dashed line is the 5-fold line, and the thick dashed line is the 10-fold line.

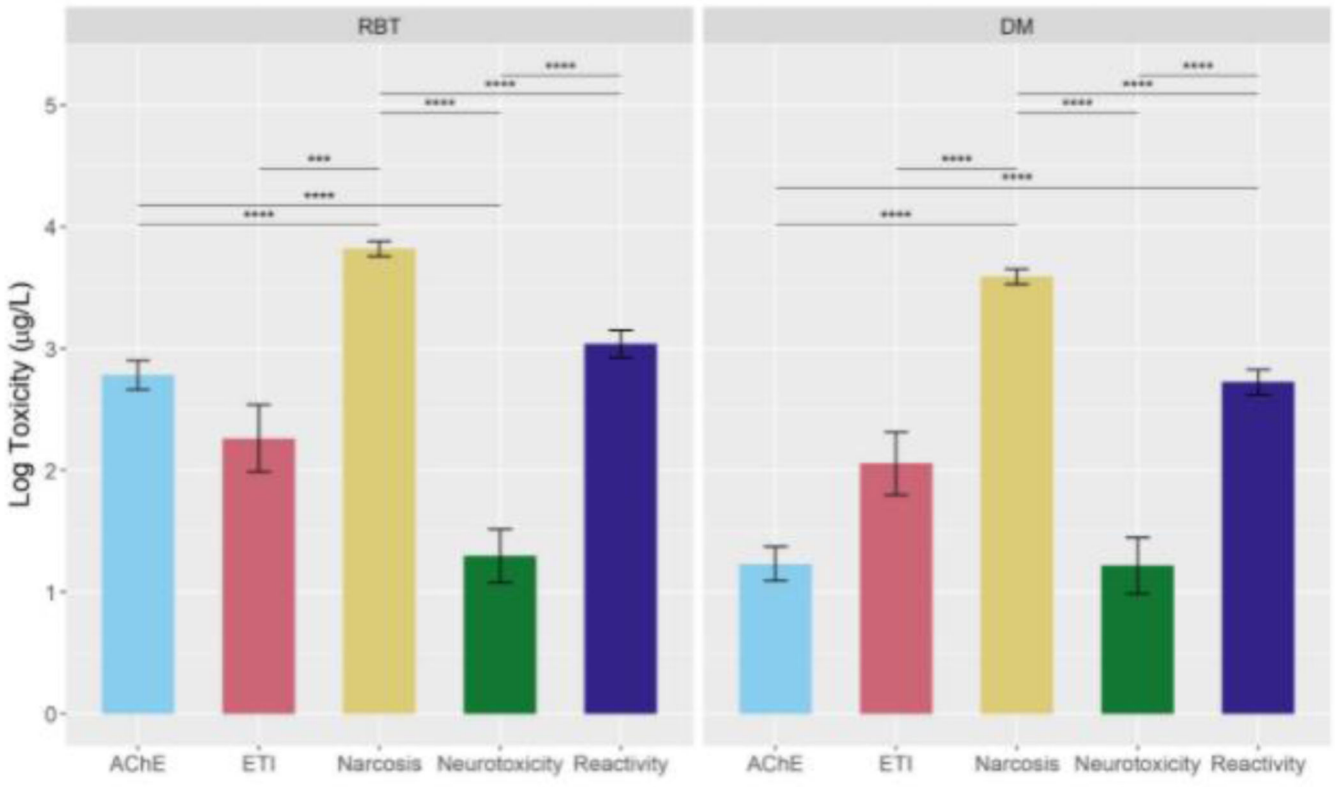

Figure 6.

Mean log toxicity among chemicals grouped by broad MOA for rainbow trout (RBT; left panel) and D. magna (DM, right panel). The mean log toxicities of broad MOA groups were compared for each species using separate one-way ANOVAs and/or pairwise t-tests. Broad MOAs that were significantly different from each other ($p < 0.05$) are identified by asterisks (*** $p < 0.001$, **** $p < 0.0001$). Abbreviations: AChE = Acetylcholinesterase inhibition; ETI = Electron transport inhibition.

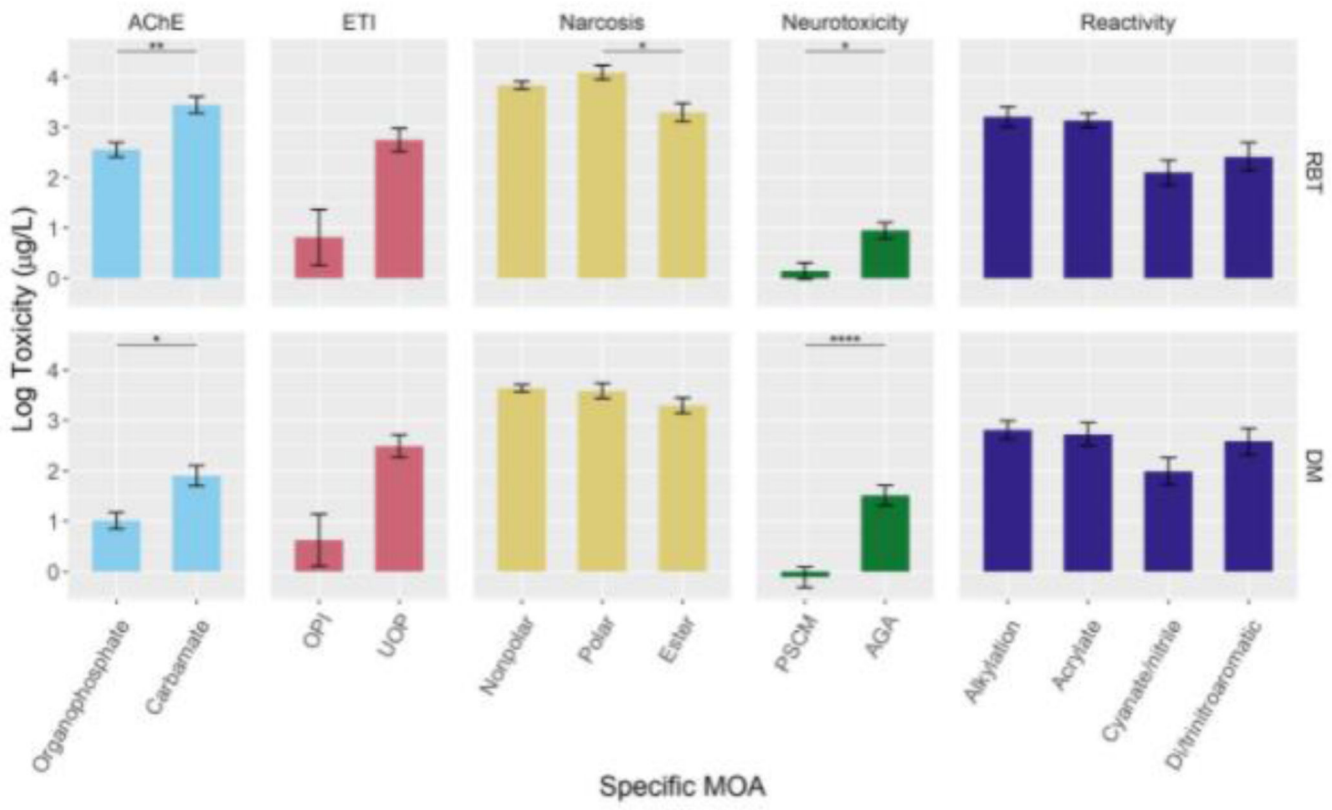

Figure 7.

Mean log toxicity of chemicals grouped by specific MOA for rainbow trout (RBT; top panels) and D. magna (DM; bottom panels). The mean log toxicity of specific MOA groups was compared within each broad MOA category, for both species, using separate one-way ANOVAs and/or pairwise t-tests. Specific MOAs that were significantly different from each other ($p < 0.05$) are identified by asterisks (* $p < 0.05$, ** $p < 0.01$, **** $p < 0.0001$). Abbreviations: OPI = oxidative phosphorylation inhibition; UOP = uncoupling oxidative phosphorylation; PSCM = pyrethroid sodium channel modulation; AGA = alicyclic GABA antagonism.

Table 1.

Breakdown of broad and specific MOAs for chemicals represented in the dataset from this study (617 compounds total). The number of chemicals in each group is also represented.

Toxicity Classification	Broad MOA	Specific MOA	No. of Chemicals
Baseline Toxicity	Narcosis *	Ester *	34
		Nonpolar *	273
		Polar *	55
		Other	7
Excess Toxicity	AChE Inhibition *	Carbamate *	19
		Organophosphate *	58
	Electron Transport Inhibition *	Oxidative phosphorylation inhibition †	7
		Uncoupling oxidative phosphorylation *	20
	Iono/Osmoregulatory/ Circulatory impairment	Anticoagulation	4
		Methemoglobinemia	1
	Narcosis *	Ester *	34
		Nonpolar *	273
		Polar *	55
		Other	7
	Neurotoxicity *	Alicyclic GABA antagonism *	7
		Diphenyl sodium channel modulation	7
		GABA agonism	2
		nAChR agonism	6
		Pyrazole GABA antagonism	1
		Pyrethroid sodium channel modulation *	22
		Sodium channel blocking	1
		Strychnine	1
		Other	3
	Reactivity *	Acrylate *	7
		Alkylation *	30
		Carbonyl	11
		Cyanate/nitrile †	8
		Di/trinitroaromatic *	13
		Hydrazine	2
		Other	12

* MOAs with significant QSAR models.

† MOAs with significant QSAR models only when using mole-based (μmol/L) data.

Table 2.

Ranges of parameters and summary statistics for log Kow-acute toxicity regression models for each mode of action. Ranges represent the differences in values across taxa models for a given MOA. Data shown is for mass-based models (μg/L) only.

Toxicity Classification	MOA	n[1]	Log Kow	Median Toxicity Value (μg/L)	R^2	Slope[2]	Log Intercept	Predictions within 5-fold	
								Broad/ Specific MOA Models	Baseline Toxicity Models
Baseline Toxicity	**Narcosis**	204–351	−5.4–7.16	3397–13958	0.311–0.581	−0.606 - -0.417	4.70–5.65	70.4%	–
	Ester	21–31	0.18–4.73	1100–23409	0.233–0.742	−0.697 - -0.385	4.46–5.61	70.3%	–
	Nonpolar	152–262	−4.7–7.13	4095–13752	0.353–0.551	−0.628 - -0.466	4.89–5.78	69.6%	–
	Polar	24–54	−1.43–5.76	49749–226270	0.171–0.855	−0.674 - -0.422	4.47–5.75	69.4%	–
Excess Toxicity	**AChE Inhibition**	62–75	−0.85–6.31	14.3–2030	0.197–0.435	−0.408 - -0.305	2.07 – 4.08	56.4%	29.4%
	Carbamate	15–19	−0.57–3.03	29.2–4652	0.258–0.353	−0.429 - -0.353	2.59–2.64	70.6%	2.9%
	Organophosphate	47–56	−0.85–6.31	6.6–1631	0.124–0.379	−0.383 - -0.241	1.76–4.07	49.4%	37.6%
	Electron Transport Inhibition	16–26	−0.25–6.79	36.9–470	0.335–0.454	−0.647 - -0.454	3.63–4.36	56.7%	32.8%
	Uncoupling oxidative phosphorylation	10–20	0.03–6.12	253.8–1090	0.426–0.656	−0.609 - -0.446	3.97–5.08	62.4%	37.6%
	Neurotoxicity	40–53	−0.85–7.16	11.0–49.60	0.528–0.597	−0.645 - -0.557	3.89–4.44	50.5%	20.4%
	Alicyclic GABA antagonism	9–11	3.66–6.50	8.03–34.60	0.419–0.419	−0.509 - -0.509	4.20–4.20	63.6%	45.5%
	Pyrethroid sodium channel modulation	16–21	2.90–7.16	0.17–5.24	0.475–0.639	−0.764 - -0.506	3.13–4.13		0.5%
	Reactivity	31–80	−1.50–5.34	230–3460	0.223–0.354	−0.415 - -0.245	3.21–4.07	56.1%	34.3%
	Acrylate	4–7	−0.120–2.80	357–9920	0.654–0.666	−0.229 - -0.227	3.72–3.73	100%	0.0%
	Alkylation	6–30	−0.859–5.22	286–6550	0.232–0.394	−0.412 - -0.328	3.44–4.16	49.0%	42.9%
	Di/trinitroaromatic	7–11	0.87–5.34	102–4.62	0.621–0.877	−0.471 - -0.390	3.81–4.36	63.9%	52.8%

[1] N denotes the total number of chemicals with toxicity values for one species.

[2] All slopes were negative.